

PTOLEMY
AND THE FOUNDATIONS OF ANCIENT
MATHEMATICAL OPTICS:
A SOURCE BASED GUIDED STUDY



A. MARK SMITH

PTOLEMY
AND THE FOUNDATIONS OF ANCIENT
MATHEMATICAL OPTICS:
A SOURCE BASED GUIDED STUDY

Transactions
of the
American Philosophical Society
Held at Philadelphia
For Promoting Useful Knowledge
Vol. 89, Pt. 3

PTOLEMY
AND THE FOUNDATIONS OF
ANCIENT MATHEMATICAL OPTICS:
A SOURCE BASED GUIDED STUDY

A. MARK SMITH

American Philosophical Society
Independence Square • Philadelphia
1999

Copyright © 1999 by the American Philosophical Society for its *Transactions* series. All rights reserved. No electronic or mechanical reproduction of any part of this publication without the express written permission of the publisher. Front cover illustration: from Friedrich Risner's *Opticae Thesaurus. Alhazeni arabis libri septem nunc primum editi. Eiusdem liber de crepusculis et nubium ascensionibus. Item Vitelonis thuringopoloni libri X.* (Basel, 1572).

**Library of Congress Cataloging-in-Publication
Data**

Smith, A. Mark.

Ptolemy and the foundations of ancient mathematical optics: a guided study / A.M. Smith.

p. cm. — (Transactions of the American Philosophical Society, ISSN 0065—9746 v. 89, Pt. 3)

Includes bibliographical references and index.

ISBN 0-87169—893—5

1. Optics—History. 2. Science, Ancient. 3. Ptolemy, 2nd cent.

I. Title. II. Series.

QC352.S58 1999

535—dc2l

99—15221

CIP

CONTENTS

Preface	7
Introduction	11
Section I: Ancient Theories of Visual Perception	23
The Physics of Vision	23
The Intromissionist Alternative	24
The Extramissionist Alternative	27
Intermediate Theories	28
The Visual Ray Theory	30
Color and Color-Perception	32
The Physiology of Vision	34
The Psychology of Visual Perception	38
The Aristotelian Account	38
Ptolemy's Account	40
Section II: Optics Proper: Analysis of Direct Vision	51
The Visual Cone	51
The Visual Perception of Physical Space	54
Binocular Vision	67
Section III: Catoptrics: Analysis of Vision by Reflected Rays	79
The Law of Equal Angles	79
Multiple Reflections and Multiple Images	83
Plane and Convex Mirrors	83
Concave Mirrors	84
The Principles of Image-Location	94
Image-Formation and Distortion	100
Plane Mirrors	101
Convex Mirrors	105
Concave Mirrors	110
Visual Effects from Composite Mirrors	119

Section IV: Dioptrics: Analysis of Vision by Deflected Rays 127

 Observation and Explanation of the Phenomenon 127

 Practical Application: The Problem
 of Atmospheric Refraction 134

 Image-Location as a Function of the
 Shape of the Refracting Surface 137

 Size-Distortion 140

Section V: Analysis of the Rainbow and of Burning Mirrors 149

 The Rainbow 149

 Burning Mirrors 153

Conclusion 161

Index 169

PREFACE

The original inspiration for this book came from Daniel Jones, Assistant Director of Interpretive Programs for the NEH, who urged me some years back to submit a proposal within the compass of a new guided-study program in the history of science. His suggestion could not have been more timely. I was then working on an English translation of Ptolemy's *Optics*, and the opportunity to use it as a springboard for a general text-based study of Greek mathematical optics was irresistible. I proposed, the NEH disposed, and this is the result.

Why this particular study at this particular time? I can think of at least two reasons. The first is historiographical. Virtually all survey-accounts of ancient science leave the mistaken impression that ancient ray-theory shared the same basic aims and methods as its modern counterpart and, therefore, that the two are genetically linked. The source of confusion lies in the fact that, for the sake of simplicity and brevity, textbook authors emphasize those aspects of Greek ray-analysis (as applied specifically to reflection and refraction) that seem most innovative and forward-looking and that are thus most resonant to the modern reader. Viewed selectively in this way, of course, Greek ray-theory looks very much like its modern counterpart. In reality, however, the two have little more in common than neolithic trepanation and modern neurosurgery. Whatever their crude similarities at a procedural level, ancient and modern ray-theory are worlds apart in conceptual and methodological foundations, and the same holds for their fundamental aims.

Take as an illustration Ptolemy's attempt to find the governing principle of refraction. The most striking thing about that attempt is its sophisticated observational basis. In a series of carefully-contrived experiments measuring refraction from air to water, air to glass, and water to glass, Ptolemy plotted r against i at ten-degree intervals from $i = 0^\circ$ to $i = 80^\circ$.¹ The resulting r -values are, on the whole, so near to what is predicted by the modern sine-law that we are left to marvel at the closeness of the call. What could be more natural than to conclude that Ptolemy failed to discover the appropriate (and exact) sine-relations not for conceptual or methodological reasons, but because his experimental results were inadequately precise? Much the same can be said of

his treatment of reflection: although Ptolemy was clearly on the right track, he went astray at critical points, particularly in his analysis of image-formation in concave mirrors. What else but imprecision or inadvertence could have led to such errors?

From a retrospective standpoint, this evaluation of Ptolemy's optical analysis and its shortcomings makes sense. After all, Ptolemy did fail to find the sine-law of refraction, and his treatment of reflection is unquestionably deficient by modern standards. But such an evaluation misses the point, for it assumes that Greek ray-theory was designed to accomplish the same ends as its modern counterpart. This assumption is as false as the interpretive structure built upon it is unilluminating. Ptolemy in fact "failed" to find the sine-law of refraction because he was looking for something else, and the "erroneous" conclusions he reached in his analysis of reflection were not erroneous at all: viewed within their proper context, they were perfectly consistent with both "fact" and theoretical expectation. The problem, at bottom, is that ancient ray-theory was intended to explain sight, not the radiation of light. If, then, Ptolemy and his fellow visual ray theorists can be said to have strayed from "the right track"—i.e., the one leading directly to modern ray theory—it is because they were embarked on an entirely different train with an entirely different destination in mind.

The second justification for this study involves the state of current scholarship. Over the past couple of decades the history of ancient optics has undergone a fundamental reappraisal in regard to both the basic sources and their interpretive context. As part of this reappraisal, several key texts have finally been rendered into English. The time is thus ripe for an introductory study that incorporates both the new wealth of translated sources and the interpretive insights that have emerged in recent years. Accordingly, what I offer here is not an anthology in the manner of Cohen and Drabkin's *A Source Book in Greek Science*² or the more recent *L'Ottica dalle origini all'inizio del '700* by Fabio Bevilacqua and Maria Grazia Ianniello.³ Such anthologies have their uses, but their intent is to illustrate a few select points, not to examine the subject in depth. My purpose is somewhat broader. Using actual sources as a foundation, I hope to show how ancient mathematical optics developed in response not only to certain theoretical imperatives but also to empirical evidence. The study that follows is therefore thesis-driven, to be sure, yet I have borne in mind throughout that, rather than force assent to my way of thinking, my primary obligation is to open a path for readers to reach their own understanding.

In assuming the role of guide rather than of scholarly apologist, I have tried to make this study accessible to the widest possible audience. For one thing, the technical demands are relatively light. All that is needed is a basic

grounding in plane (Euclidean) geometry. For another thing, although the ostensible focus of the study is on ray-analysis, considerable attention is paid throughout to issues of perceptual psychology, physiology, and even methodology. So there is plenty of grist for the mill of anyone not inclined to work punctiliously through the mathematical demonstrations. Practically speaking, then, this study is geared to readers with a fairly wide range of skills and interests, from the willing undergraduate to the academic professional, from the philosophically- to the technically-minded: to anyone, in short, seeking a source-based introduction to optics and visual theory as it developed in antiquity.

Now to acknowledgments. Let me start by expressing my deep gratitude to the National Endowment for the Humanities for their generous support in the completion of this project. Thanks are also due to my home institution, the University of Missouri, for unstinting supplementary support. For inspiration and insight, given either directly or indirectly, the following scholars deserve special mention: Albert Lejeune, whose pioneering studies of Euclidean-Ptolemaic optics helped me see things I would otherwise surely have missed; David C. Lindberg, who guided me early on as I groped my way toward an appreciation of ancient and medieval optics; and Gérard Simon, whose clarity of thought about Euclidean-Ptolemaic optics aided me immeasurably in focusing my own. For various editorial criticisms and suggestions, finally, I want to thank Lois L. Huneycutt, Alan Bowen, and Jennifer Erickson, the first for her incisive critique of my ideas and their framing and the latter two for their punctilio in weeding out errors and infelicities of expression. Whatever faults remain are due either to my inattention to detail or my refusal to yield to better judgment.

NOTES

¹The symbol i represents the angle of incidence as measured according to the perpendicular dropped to the refracting surface at the point of refraction; r is the angle of refraction as measured according to that same perpendicular. For a diagram illustrating these two angular relationships, see figure 4.2 on p. 130 below.

²Cambridge: Harvard University Press, 1948.

³Turin: Loescher Editore, 1982.

INTRODUCTION

One of the most remarkable achievements of Greek science was the invention of the ray-concept and the attendant development of mathematical optics on its basis. Precisely when this simple yet ingenious analytic expedient was first applied to the study of sight and light is a matter of speculation, but we can trace its origins to the mid-fourth century B.C. at latest, and probably even earlier. In the *Physics*, for example, Aristotle included optics among “the more natural of the branches of mathematics,” adding by way of elucidation that, “while geometry investigates natural lines but not *qua* natural, optics investigates mathematical lines, but *qua* natural, not *qua* mathematical.”¹ Aristotle’s grasp of mathematical optics was not just theoretical; in the third book of the *Meteorology* he actually attempted to explain the formation of haloes and rainbows on the basis of ray-analysis, ultimately imputing the phenomenon to reflection.²

As might be expected, ray-theory underwent considerable development from the time of Aristotle to the appearance of Ptolemy’s *Optics* in the second half of the second century A.D., when ancient mathematical optics reached its culmination in terms of both analytic scope and theoretical specification. It is our purpose in this study to trace the course of that development, using as a textual core four sources specifically devoted to mathematical optics. Composed around 300 B.C., Euclid’s *Optics* represents the earliest of these sources. In it the basic theoretical framework for all subsequent development is laid. Next comes the Pseudo-Euclidean *Catoptrics*, which forms a natural companion-piece to the *Optics* but may reflect a somewhat later stage of development. Dating from around the mid-first century A.D., our third source, Hero of Alexandria’s *Catoptrics*, offers important insights into the evolution of post-Euclidean optics. And finally, rounding out our list is Ptolemy’s *Optics*, which was written sometime between 160 and 170 A.D. as a comprehensive synthesis in the mold of its astronomical counterpart, the *Almagest*.³

In addition to this somewhat meager core of technical sources, we will also draw upon certain supplementary works. Among these are: Plato’s *Timaeus* (mid-fourth century B.C.), Aristotle’s *Meteorology*, *On the Soul*, and *Sense and Sensibilia* (mid-fourth century B.C.), the Pseudo-Aristotelian *Problems* (fourth

century B.C.?), Theophrastus' *On the Senses* (late fourth century B.C.), Diocles' *On Burning Mirrors* (early second century B.C.), Lucretius' *On the Nature of Things* (mid-first century B.C.), and Galen's *On the Usefulness of the Parts of the Body* and *On the Opinions of Plato and Hippocrates* (later second century A.D.).⁴ The reason for including these supplementary sources is to broaden the context of our study, particularly in regard to issues of visual perception and its physical, physiological, and psychological basis.

At first glance, this emphasis on visual perception might seem misplaced in a study of mathematical optics. But it is crucial to realize that, unlike its modern counterpart, ancient mathematical optics was concerned not with the physics of light but with the process of sight. Indeed, the very term "optics" (Greek: *optica*, from *ops* = "eye") indicates that, for ancient optical theorists, the cardinal reference-point was the eye. The ray was therefore conceived of as a line of sight rather than a trajectory for light, its fundamental purpose being to account for the visual perception of spatial reality. On that basis, ancient mathematical optics developed according to a tripartite analytic structure. Optics proper, on the one hand, was concerned with the appearances resulting from an unimpeded line of sight between eye and object. Catoptrics, on the other, was concerned with the appearances resulting from a fully broken line of sight. Its primary aim was therefore to explain mirror-images and their deviation from the objective reality they represent. Dioptrics, finally, dealt with the appearances resulting from a partially broken line of sight. In this case, too, the primary aim was to explain images and their deviation from objective reality.

The study that follows reflects this basic analytic structure. It consists of five sections, the first of which focuses upon the development of visual theory from roughly the time of Aristotle to that of Ptolemy. We will start this section by examining various physical explanations of sight proposed in antiquity, paying particular attention to theories of light and color. We will then turn to the physiology of vision, taking a brief look at the development of ocular anatomy in relation to theories of sensation and the nervous system. That done, we will close with a look at the psychology of visual perception as understood, primarily, by Aristotle and Ptolemy.

Taking the foregoing analysis as context, we will focus in section two upon ray-theory itself, our primary goal being to understand precisely what the ray was thought to represent in Greek optics and how it was used to account for the visual perception of space. We will then turn to catoptrics in section three, examining not only theoretical but also empirical and mathematical approaches to reflection from plane, convex, and concave mirrors. Section four will be devoted to dioptrics, our focus in this case being upon Ptolemy's analysis of refraction through various media and his subsequent attempt to derive a gen-

eral principle relating angles of incidence and refraction. The fifth and final section will be devoted to early analyses of the rainbow and burning mirrors, two topics that seem not to have been integral to the mainstream study of optics. To put this text-based study into proper historical context, finally, we will conclude with a cursory look at the fate of the Euclidean-Ptolemaic optical tradition during late antiquity and the Middle Ages. Suffice it to say, our purpose in all this is to give a mere sketch, not a detailed picture.

As far as format is concerned, the study will consist of excerpts drawn from the sources mentioned above. Those excerpts will be tied together by short introductory or explanatory passages intended to provide a coherent framework for them. The excerpts themselves will be numbered consecutively within their respective sections, and each excerpt will be given a full source-reference. Whenever necessary, intra-textual annotation will be provided in order to clarify points. In most cases, I have relied on published translations for the source-texts, but I have supplied my own translations in the case of Euclid's *Optics*, the Pseudo-Euclidean *Catoptrics*, and Hero's *Catoptrics*. Since this study is introductory in nature, I have, as far as possible, ignored scholarly disputes about the precise meaning and intention of various textual passages that are open to debate. For the sake of simplicity and clarity, therefore, I have at times taken interpretive stands that are more definite (some might even say more reckless) than the cited texts may actually warrant, particularly in the first section dealing with theories of visual perception. Part of the problem, of course, has to do with the texts themselves. After all, no text simply speaks for itself, and the farther removed in time and culture a text is, the more garbled its message is likely to seem. But a considerable part of the interpretive problem has to do with the type and quality of textual evidence available, a point that merits some amplification.

The Problem of Sources: There is no shortage of modern literature devoted either wholly or in part to the development of Greek optics. Nor is there any shortage within this literature of confident assertions about the course of that development, its *dramatis personae*, its theoretical underpinnings, or its ulterior significance. When we move from these secondary studies to the actual sources upon which they are based, however, we cannot help but be struck by the extreme tenuousness of that basis. Rarely, in fact, has so much been made by so many of so little. A brief discussion of the textual basis for our current understanding of Greek optics is therefore in order.

Drawing upon an earlier compilation by Thrasyllus (fl. early first century A.D.), Diogenes Laertius (fl. early third century A.D.?) lists some seventy treatises thought to have been written by Socrates' notorious contemporary, Democritus (460-357 B.C.). If we are to judge by the titles, no fewer than

four of the listed works were devoted entirely or in significant part to optical matters. Yet not one of the four has survived. In fact, the entire *oeuvre* of Democritus is now lost, and there is no telling how much of it survived even up to the time of Diogenes Laertius. Likewise, among the almost 230 works attributed by the same author to Aristotle's disciple, Theophrastus (fl. c. 320 B.C.), at least four were concerned with optical subjects. Of these four, the only survivor is *On the Senses*, in which the treatment of vision forms only part, albeit the major part, of a more general study of sensation. Not an exposition of Theophrastus' own theory of sensation, however, this treatise is actually a critical commentary on the theories of certain of his predecessors, Democritus included.

These two cases illustrate most, if not all, of the basic problems facing the historian of optics from Presocratic times up to, and well beyond, the late Hellenistic period. The first and most obvious of these problems is the dearth of primary texts available to us today. We know, for instance, that, as part and parcel of a general interest in natural philosophy, the study of optics began in earnest long before Aristotle. We also know that this initial phase of study spawned numerous treatises on light, color, and vision. Both Theophrastus and Diogenes Laertius, among others, attest to that fact. Against all odds, however, every one of these seminal works has sunk into oblivion. Apart from scattered passages in Plato's dialogues, the earliest extant optical texts are to be found in the Aristotelian corpus. Almost all direct traces of the development of optical thought before Aristotle have thus been effectively erased.

Even for the period between Aristotle and Ptolemy we have almost nothing in the way of primary sources, and the little we do have is chronologically skewed, most of it—with the signal exception of Theophrastus' *On the Senses* and Euclid's *Optics*—dating from after 100 B.C. Hence, just as for the period up to roughly 350 B.C., so for the period from roughly 300 B.C. to 100 B.C., our understanding of the development of optical thought is hampered by a lack of direct textual evidence.

Failing primary sources, then, we are forced to rely heavily upon secondary or indirect sources. Scarcity is not the problem here. On the contrary, we have a rather wide variety of such sources, ranging from critical commentaries (e.g., Theophrastus' *On the Senses*) to doxographical accounts (e.g., Diogenes Laertius' *Lives*). Even primary sources can double as secondary ones. Aristotle, for instance, supplies a good deal of information about previous theories of vision in order to lay the background for his own discussion. Galen, too, tells us a great deal about Stoic visual theory in the course of outlining and defending his own. Like the indirect sources themselves, moreover, the information they convey ranges in kind from direct quotations ("X said 'Y'"), through paraphrases ("X said that Y"), to oblique citations ("X is said to have

claimed that Y").

Obviously, what is at issue here is reliability. How much credence can we actually give to these sources and the information they impart? The answer is contingent upon a number of factors, among which time looms especially large. For example, both Aristotle and Diogenes Laertius have much to tell us about Democritus, yet Aristotle was almost contemporaneous with Democritus, while Diogenes Laertius lived at least seven centuries after him. Moreover, Diogenes was prone to give oblique citations, which indicates that he frequently drew his information at second hand without much critical concern. Ought we then to give more credence to Aristotle? Not necessarily, because Aristotle was notoriously selective in his use and interpretation of sources, sometimes willfully distorting them in order to reinforce the superiority of his own positions.

A further problem, which is particularly germane in the case of the four primary sources upon which this present study is based, is that of authenticity. As is so often the case with ancient texts, neither the authenticity nor provenance of those sources can be established with absolute certainty. There are several reasons. First, none of the extant manuscripts containing them dates from before the Middle Ages, so the texts as we know them are products of successive redactions that may well have altered the original beyond all recognition. Second, only two of the texts have reached us in Greek. The remaining two are available only in Latin translation. Thus, whatever textual deviations might naturally accrue through successive redactions can only have been compounded by such translation. Finally, there is the problem of false attribution. Whether purposely or not, many ancient works have been misascribed in the course of textual transmission. At least one of our key texts, Hero's *Catoptrics*, was long attributed to the wrong author (Ptolemy), and the same could well hold for others.

While such problems may prevent us from achieving absolute certainty, they need not keep us from establishing the authenticity of our sources with at least relative assurance. For instance, there is a considerable body of testimony from antiquity and the early Middle Ages confirming that optical works by Euclid, Hero, and Ptolemy were in circulation during that period. In the case of Hero's *Catoptrics*, in fact, it was just such testimony that permitted the modern editor to determine its true authorship in spite of its longstanding misattribution to Ptolemy. Moreover, a comparison of the works among themselves may provide at least a rough sense of chronological order. There can be no doubt, for example, that Ptolemy's *Optics* represents a more advanced stage (at least conceptually) in the development of ray-theory than Euclid's *Optics* and *Catoptrics*, or for that matter than Hero's *Catoptrics*, which in turn represents a clear advance over its Euclidean counterpart. Granted, in actual chro-

nological rather than in ideal evolutionary terms, apparent progenitor may well turn out to be offspring, but in the case of our four treatises all evidence points to a parallel between evolutionary and chronological succession. Thus, however problematic the precise chronology and attribution of our source-texts may be, we have at least some assurance that they mark successive points along a line of theoretical development. Following is a list, with brief descriptions, of the basic sources used in this study.

1. *The Four Primary Texts*

Euclid's *Optics*: Best-known for the *Elements*, his magisterial compendium of plane and solid geometry, Euclid apparently taught at Alexandria in the late fourth and early third centuries B.C., when that city began to rival, if not surpass, Athens as an intellectual and cultural center within the Hellenistic world. One of his relatively minor works, the *Optics*, has come down to us in Greek, the earliest manuscript dating from the tenth century A.D. In fact, as the great Danish philologist J. L. Heiberg discovered, there are two distinct versions of the *Optics*. One of them is thought to represent a recension by Theon of Alexandria (later fourth century A.D.), whereas, according to Heiberg's analysis, the other is genuinely Euclidean. Critically edited with facing Latin translations, both versions were published by Heiberg in 1895. The two have since been translated into French by Paul Ver Eecke. The full text of the supposedly genuine *Optics* has been translated by H. E. Burton, and excerpts have been translated and published in M.R. Cohen and I. E. Drabkin's *Source Book in Greek Science*. Although Heiberg's categorization of the two versions of Euclid's *Optics* has gone unchallenged up to now, one scholar has recently questioned it, suggesting instead that the putatively genuine version may actually be the later of the two. Whatever the case, however, there is no dispute that the propositional content of both versions is fundamentally Euclidean.⁵

Hero of Alexandria's *Catoptrics*: There has been considerable disagreement over the years about Hero of Alexandria's *floruit*, but the current consensus is that it belongs somewhere within the middle decades of the first century A.D. Like the *Mechanics* and *Pneumatics*, Hero's *Catoptrics* reveals a strong practical bent, much of the treatise being devoted to the arrangement of mirrors for particular visual effects. Unfortunately, the version that we currently possess exists only in a Latin translation produced by William of Moerbeke in 1269; the Greek original has yet to be found. Worse yet, this particular version appears to be an abridgment of the original, consisting of snippets strung together with little or no regard to logical sequence. Falsely ascribed to Ptolemy since the Middle Ages, the *Catoptrics* has been properly credited to Hero since

the very turn of this century, when it was critically edited and published with a facing German translation by Wilhelm Schmidt. English excerpts can also be found in Cohen and Drabkin's *Source Book*.⁶

The “Pseudo-Euclidean” *Catoptrics*: Although this work is consistently ascribed to Euclid in the extant manuscripts, its modern editor, J. L. Heiberg, concluded for stylistic reasons that it could not be genuinely Euclidean, claiming, instead, that it was a late compilation by Theon of Alexandria containing both Euclidean and post-Euclidean elements. On that understanding, he published the critical text along with facing Latin translation in tandem with the two versions of Euclid's *Optics*. A French translation was subsequently published by Ver Eecke along with his translation of the *Optics*. Following Heiberg's suggestion, the respected Belgian scholar, Albert Lejeune, claimed to have discerned three basic chronological strata in the work, the earliest dating to the period between Euclid and Archimedes, the next to around the time of Hero, and the last to the time of Theon. According to Lejeune's account, more than 70% of the theorems in the work belong to the two earlier strata; even more to the point, nearly 80% of that total belongs to the earliest stratum. Recently, however, a persuasive argument has been made against both Heiberg and Lejeune in defense of the work's authenticity.⁷ I will therefore treat this work as genuinely Euclidean throughout the remainder of this study.

Ptolemy's *Optics*: Born around 100 A.D., Ptolemy passed his working life in or near Alexandria. Before his death in the late 170's he had written a wide variety of works that included four major compendia: the *Syntaxis Mathematike* (or *Almagest*, as it came to be known among the Arabs), the *Geography*, the *Tetrabiblos*, and the *Optics*, which was probably written in the later years of his life. Translated into Arabic by the early tenth century at latest, this work has left no trace in either Greek or Arabic. All that survives is a badly mangled twelfth-century Latin version translated from the Arabic by a Byzantine functionary (the “Emir” Eugene) in Norman Sicily. In addition to various minor lacunae and textual inconsistencies, this version lacks the entire first book and ends abruptly in the middle of the fifth. Despite some controversy over the work's authenticity, there is no compelling reason to doubt that it is genuinely Ptolemaic. A critical Latin edition was produced by Albert Lejeune in 1956, and his French translation of the text has since been published posthumously. The excerpts used in this study have been drawn from my English translation of 1996.⁸

2. *Supplementary Sources*

Aristotle's *Meteorology*, *On the Soul*, and *Sense and Sensibilia*:⁹ A student of Plato, Aristotle (384-322 B.C.) founded his own school, the Lyceum, after

Plato's death. His account of the rainbow in the *Meteorology* provided the standard explanation until the later Middle Ages. *On the Soul* represents Aristotle's most extensive treatment of the psychological and epistemological entailments of sense-perception. Part of the so-called *Parva naturalia*, Aristotle's *Sense and Sensibilia* offers a more fundamental examination of sense-perception and its physical ramifications.

Diocles' *On Burning Mirrors*:¹⁰ We know virtually nothing about Diocles beyond his probable *floruit* of 190-180 B.C. Surviving only in Arabic translation, his *On Burning Mirrors* contains the earliest known treatment of the focusing properties of parabolic sections.

Galen's *On the Usefulness of the Parts of the Body* and *On the Opinions of Hippocrates and Plato*:¹¹ Remarkable for his skills as both anatomist and medical theoretician, Ptolemy's younger contemporary, Galen (c. 130-c. 200 A.D.), was deeply indebted to the Alexandrian medical tradition. Despite his vehement attacks against the Stoics, he was heavily influenced by them, particularly in his physical leanings. A compendious description of the various organs of the body, Galen's *On the Usefulness of the Parts* contains a long discussion of the eye and its functioning in the tenth book. In *On the Opinions of Hippocrates and Plato* Galen attacks various doctrines of the Stoic philosopher, Chrysippus, and in the process gives us valuable insight into those doctrines and their meaning. The work therefore represents a crucial source for our understanding of the Stoic theory of vision--albeit refracted through a distinctly Galenic prism.

Lucretius' *On the Nature of Things*:¹² By the author's own testimony, this Epicurean manifesto was intended to free its readers from the trammels of religious superstition and the fears inspired by it. In it, Lucretius (c. 99-c. 55 B.C.) attempts to explain the full range of natural phenomena on the basis of atoms and their more-or-less free associations. Accordingly, his account of vision in the fourth book of the treatise reduces the process to atomic actions and reactions.

Plato's *Timaeus*:¹³ Founder of the Academy at Athens, Plato (427-348/47 B.C.) spent much of his later life shuttling between Athens and Magna Graecia, where he is supposed to have come under Pythagorean influence. Among the very latest of his dialogues, the *Timaeus* certainly reveals traces of such influence. As part of the cosmogonic/cosmological account offered in that work, Plato includes a relatively extensive discussion of sensation, part of which deals with vision.

The Pseudo-Aristotelian *On Colours* and *Problems*:¹⁴ The first of these works offers a fundamentally Peripatetic discussion of color and color-perception,

discussing the basic formation of colors and the ambient circumstances that can modify our perception of them. The *Problems* is an extensive compilation that touches on a remarkable range of subjects, including visual theory. Built around an early Peripatetic core, the work as a whole appears to be the result of constant accretions over several centuries.

Theophrastus' *On the Senses*:¹⁵ A close disciple of Aristotle and his successor to the headship of the Lyceum in 322, Theophrastus (c. 372-c. 287 B.C.) was evidently a prolific author, being credited with over 200 works. Unfortunately, only a handful has survived, among which *On the Senses* represents an extensive critique of various contemporary accounts of vision.

EDITIONS AND TRANSLATIONS USED IN THIS STUDY

Aristotle, *Meteorology*, trans. E. W. Webster, in Jonathan Barnes, ed., *The Complete Works of Aristotle* (Princeton: Princeton University Press, 1984), vol.1, pp. 555-625.

_____, *On the Soul*, trans. J. A. Smith, in Jonathan Barnes, ed., *The Complete Works of Aristotle* (Princeton: Princeton University Press, 1984), vol.1, pp. 641-92.

_____, *Sense and Sensibilia*, trans. J. I. Beare, in Jonathan Barnes, ed., *The Complete Works of Aristotle* (Princeton: Princeton University Press, 1984), vol. 1, pp. 693-713.

Diocles, *On Burning Mirrors*, ed. and trans. G. J. Toomer, *On Burning Mirrors: The Arabic Translation of the Lost Greek Original* (Berlin/Heidelberg/New York: Springer-Verlag, 1975).

Euclid, *Catoptrics*, ed. J. L. Heiberg, *Euclidis opera omnia*, vol. 7 (Leipzig: Teubner, 1895), pp. 285-343.

_____, *Optics*, ed. J. L. Heiberg, *Euclidis opera omnia*, vol. 7 (Leipzig: Teubner, 1895), pp. 1-121.

Galen, *On the Opinions of Hippocrates and Plato*, trans. Phillip De Lacy, *Corpus Medicorum Graecorum*, V, 4, 1, 1-2 (Berlin: Akademie Verlag, 1980).

_____, *On the Usefulness of the Parts*, trans. Margaret Talmadge May, *Galen on the Usefulness of the Parts of the Body*, 2 vols. (Ithaca, N.Y.: Cornell University Press, 1968).

Hero of Alexandria, *Catoptrics*, ed. and trans. Wilhelm Schmidt, *Heronis Alexandrini opera quae supersunt omnia*, vol. 2:1 (Leipzig: Teubner, 1900), pp. 301-373.

Lucretius, *On the Nature of Things*, trans. W. H. D. Rouse and M. F. Smith, *Lucretius: De rerum natura*, Loeb Classical Library, 2nd Edition (Cambridge: Harvard University Press, 1982).

Plato, *Timaeus*, trans. Francis M. Cornford, *Plato's Cosmology: The Timaeus of Plato Translated with a Running Commentary* (New York: Liberal Arts Press, 1957).

Pseudo-Aristotle, *On Colours*, trans. T. Loveday and E. S. Forster, in Jonathan Barnes, ed., *The Complete Works of Aristotle* (Princeton: Princeton University Press, 1984), vol. 1, pp. 1219-1228.

———, *Problems*, trans. E. S. Forster, in Jonathan Barnes, ed., *The Complete Works of Aristotle* (Princeton: Princeton University Press, 1984), vol. 2, pp. 1319-1527.

Ptolemy, *Optics*, trans. A. Mark Smith, *Ptolemy's Theory of Visual Perception: An English Translation of the Optics with Introduction and Commentary*, Transactions of the American Philosophical Society 86.2 (Philadelphia: American Philosophical Society, 1996).

Theophrastus, *On the Senses*, trans. George M. Stratton, *Theophrastus and the Greek Physiological Psychology before Aristotle* (London: Allen and Unwin, 1917).

NOTES

¹*Physics*, II, 2, 194a7-12, trans. R. P. Hardie and R. K. Gaye, in Jonathan Barnes, ed., *The Complete Works of Aristotle: The Revised Oxford Translation* (Princeton: Princeton University Press, 1984), vol. 1, p. 331.

²*Meteorology*, III, 2-6; see section V, 1, pp. 149-153 below for Aristotle's explanation of the rainbow.

³For further details on these primary sources and their textual history, see pp. 16-17 above.

⁴For further details on these supplementary sources, see pp. 17-19 above.

⁵The full citation for Heiberg's edition can be found on p. 19 above. For the critique of Heiberg's classification of the two versions, see Wilbur Knorr, "Pseudo-Euclidean Reflections in Ancient Optics: A Re-examination of Textual Issues Pertaining to the Euclidean *Optica* and *Catoptrica*," *Physis* 31 (1994): 1-45. Ver Eecke's French translation of both versions can be found in *Euclide: L'Optique et la Catoptrique* (Paris: Blanchard, 1959), pp. 1-98. For the English translation, see H. E. Burton, "The Optics of Euclid," *Journal of the Optical Society of America*, 35 (1945): 357-72. For the English excerpts, see Cohen and Drabkin, *Source Book*, pp. 257-61.

⁶The full citation for Schmidt's edition can be found on p. 20 above. For the English excerpts, see Cohen and Drabkin, *Source Book*, pp. 261-68.

⁷The full citation for Heiberg's edition can be found on p. 19 above. Ver Eecke's French translation can be found in *Euclide: L'Optique et la Catoptrique*, pp. 99-123. For Lejeune's analysis of the theorematic content of the *Catoptrics*, see *Recherches sur la catoptrique grecque d'après les sources antiques et médiévales*, in *Mémoires de l'Académie Royale de Belgique: Classe des Lettres et des Sciences Morales et Politique*, vol. 52:2 (Brussels: Palais des Académies, 1957), pp. 112-36, esp. 112-13. For the argument against Heiberg and Lejeune, see Wilbur Knorr, "Pseudo-Euclidean Reflections" and "Archimedes and the Pseudo-Euclidean *Catoptrics*: Early Stages in the Ancient Geometric Theory of Mirrors," *Archives internationales d'histoire des sciences* 35 (1985): 28-105. For further recent corroboration of Knorr's position, see Ken'ichi Takahashi, *The Medieval Latin Traditions of Euclid's Catoptrica* (Fukuoka: Kyushu University Press, 1992), pp. 13-37, and Gérard Simon, "Aux origines de la théorie des miroirs: sur l'authenticité de la *Catoptrique* d'Euclide," *Revue d'histoire des sciences* 47 (1994): 259-72.

⁸The full citation for my translation can be found on p. 20 above. The recent French translation can be found in Albert Lejeune, ed. and trans., *L'Optique de Claude Ptolémée dans la version latine d'après l'arabe de l'émir Eugène de Sicile: Édition critique et exégétique augmentée d'une traduction française et de compléments* (Leyden/New York/Copenhagen/Cologne: Brill, 1989). For the argument against Ptolemy's authorship of the *Optics*, see Knorr, "Archimedes and the Pseudo-Euclidean *Catoptrics*," pp. 96-104; for the counter-argument in favor, see Lejeune, *L'Optique de Claude Ptolémée*, pp. 133*-138*.

⁹See p.19 above for full citations.

¹⁰See p. 19 above for full citation.

¹¹See p. 19 above for full citations.

¹²See p. 20 above for full citation.

¹³See p. 20 above for full citation.

¹⁴See p. 20 above for full citation.

¹⁵See p. 20 above for full citation.

SECTION I

ANCIENT THEORIES OF VISUAL PERCEPTION

Between the roughly six centuries separating the Presocratics from Ptolemy, a fairly wide range of visual theories was proposed and debated. Perhaps the easiest way to understand these theories is by topical foci. Accordingly, the following discussion will be divided into three sub-sections. In the first sub-section we will look at how various theorists explained the visual process in terms of the physical interaction between eye and visible object. In the next sub-section we will focus on the physiological explanation of sight, examining how the anatomical and physiological structure of the eye was understood to contribute to the visual process. In the third sub-section, finally, we will look at how certain thinkers explained the psychology of perception according to levels of visual and mental processing.

1. The Physics of Vision

Whatever their differences in detail, ancient Greek theories of vision all found common ground in the assumption that sight cannot occur without some physical mediation between the eye and visible objects. That belief, in turn, was rooted in a critical supposition about physical causation and sustained by an equally critical fact. The critical supposition is that material bodies, such as eyes and visible objects, can only affect one another through physical contact. In short, action at a distance is impossible. The critical fact is that, in order for vision to occur at all, there must be some spatial separation between eye and object. Direct contact simply will not do. Having, therefore, precluded both action at a distance and direct physical contact, Greek visual theorists were forced to adduce some *tertium quid* that would physically link eye and object while, at the same time, serving as a perceptual bridge between them.

Agreement over the need for such a mediating link still left ample room for debate over how this link was established. From what we know, virtually all of the theoretical responses that emerged in the course of that debate fell between two logical extremes. At one end lay intromissionism, which ex-

plains vision in terms of a physical influx passing from the visible object to the eye. It was upon a fundamentally intromissionist basis that the ancient atomists, as well as Aristotle, constructed their accounts of vision. At the other end lay extramissionism, according to which vision is explained in terms of an emission outward from the eye to visible objects. It was, of course, upon a fundamentally extramissionist basis that the visual-ray theorists from Euclid to Ptolemy constructed their accounts of vision. Finally, between the intromissionist and extramissionist extremes lay various intermediate theories predicated on the need for both influx from external sources and efflux from the eye. Among those who took this intermediate position, Plato and the Stoics figure most prominently. Let us briefly examine the three basic alternatives in turn, starting with intromissionism.

The Intromissionist Alternative: The atomist theory of vision was rooted in the assumption that physical objects continually emit atom-thick replicas of themselves. This assumption is clear in the reported claim by Epicurus (fl. c. 300 B.C.) that “particles are continually streaming off from the surface of bodies And those given off for a long time retain the position and arrangement which their particles had when they formed part of the solid bodies.”¹ Formal representations of their generating objects, these replicas (*eidola*, in Greek) serve as images for the eye (and thence the mind) into which they are transmitted. Thus, as Lucretius (fl. c. 80 B.C.) puts it in *De rerum natura*, 4.54 and 4.239 (pp. 281 and 295-97):

[1.1] Since amongst visible things many throw off bodies, sometimes loosely diffused abroad, as wood throws off smoke and fire heat, sometimes more close-knit and condensed, as often when cicadas drop their neat coats in summer, and when calves at birth throw off the caul from their outermost surfaces, . . . a thin image must also be thrown off from things, from the outermost surface of things, . . . which can be cast off in the same arrangement they were in before, preserving the shape of the object, and far more quickly Now the images of things I speak of are being carried about and thrown off scattered abroad in all directions; but because it is only with our eyes we can perceive them, therefore it happens that where we turn our sight, there all things strike upon it with shape and colour.

Since vision depends upon such material representations, or *simulacra*, Lucretius concludes in *De rerum natura*, 4.230-238 (p. 295), that sight operates in much the same way as touch:

[1.2] Since a shape handled in the dark is recognized to be the same which is seen in the clear light by day, it must be that touch and sight are moved by a like cause. Now, therefore, if we take hold of something square and it excites our feeling in the dark, in the light what square thing can fall upon

our vision, if not an image of it? Therefore there is seen to be in images a cause of vision, and without these nothing can be seen.

While both Epicurus and Lucretius seem to think that the material replicas of visible objects are the sole mediating entities in visual perception, their predecessor, Democritus (fl. c. 430 B.C.), offers a somewhat different account. Theophrastus describes that account in *On the Senses*, 50, pp. 108-111:

[1.3] Vision [Democritus] explains by the reflection (*emphasis*) [in the eye] of which he gives a unique account. For the reflection does not arise immediately in the pupil. On the contrary, the air between the eye and the object of sight is compressed by the object and the visual organ, and thus becomes imprinted (*typousthai*); since there is always an effluence of some kind arising from everything. Thereupon, this imprinted air, because it is solid and is of a hue contrasting [with the pupil²], is reflected in the eyes, which are moist. [And] Democritus himself, in illustrating the character of the “impression”, says that “it is as if one were to take a mould in wax”.

The Greek term *emphasis*, which is translated by “reflection” in this passage, can also mean “presentation” or even “image,” in which case the sense of the passage is that vision is caused by physical impressions created by the *eidola* streaming from the object. Transmitted mechanically through the air, these impressions strike the cornea, creating a formal representation in it of the original object-source. In short, vision is ultimately due to a sort of stamping-process, the result of which can be seen in the images reflected at the cornea’s surface. Since vision requires the mediation of air, what is seen is not a material replica of, but a mechanical effect created by, the visible object.

Like Democritus, Aristotle predicates his account of vision on mediation, but unlike him, Aristotle denies that vision is in any way analogous to touch. “Democritus and most of the natural philosophers who treat of sense-perception proceed quite irrationally,” Aristotle remarks in *Sense and Sensibilia*, 4, 442a29-b3 (p. 702), “for they represent all objects of sense as objects of touch. Yet, if this is really so, it clearly follows that each of the other senses is a mode of touch; but one can see at a glance that this is impossible.” Aristotle’s account of sight depends upon three preconditions. First, the visible object must have color, which is the proper object of sight and, therefore, the only thing that is *per se* visible. Second, eye and visible object must be physically linked by a continuous transparent medium, such as air or water, for without such media, color could not be transmitted from object to eye. Finally, because such media by their nature are only *potentially* transparent, they must be rendered *actually* transparent by light. As Aristotle sums it up in *On the Soul*, II, 7, 418a29-419a20 (pp. 666-67) :

[1.4] Whatever is visible is colour and colour is what lies upon what is in its own nature visible; . . . every colour has in it the power to set in movement what is actually transparent; [and] by 'transparent' I mean what is visible, and yet not visible in itself, but rather owing its visibility to the colour *of something else*; of this character are air, water, and many solid bodies. Neither air nor water is transparent because it is air or water; they are transparent because each of them has contained in it a certain substance [of which] light is the activity—the activity of what is transparent so far forth as it has in it the determinate power of becoming transparent; where this power is present, there is also the potentiality of the contrary, viz darkness. Light is as it were the proper colour of what is transparent and exists whenever the potentially transparent is excited to actuality by the influence of fire . . . The following experiment makes the necessity of a medium clear. If what has colour is placed in immediate contact with the eye, it cannot be seen. Colour sets in movement not the sense organ but what is transparent, e.g. the air, and that, extending continuously from the object to the organ, sets the latter in movement . . . [Therefore] seeing is due to an affection or change of what has the perceptive faculty [i.e., the eye], and it cannot be affected by the seen colour itself; it remains that it must be affected by what comes between. Hence it is indispensable that there be *something* in between—if there were nothing, so far from seeing with greater distinctness, we should see nothing at all.

Several features of this account are worth noting. For a start, if color is all that is properly seen, then shape, size, relative configuration, etc.—in short, the sum of visible properties for the atomists—must not be *per se* visible. Light is not *per se* visible either; its function is to actualize transparency, and it is only through fulfillment of that function that its existence can be inferred. What we vulgarly refer to as “light” is just the bright color of fiery or shining objects, not—as the atomists would have it—a material emanation diffused outward from certain kinds of bodies, such as the sun or stars. Finally, without an intervening transparent medium through which the color-impression can be transmitted, vision is impossible.

Color, however, is a inherent quality of physical objects, not a material efflux, so the transmission of color through a transparent medium to the eye entails an immediate (i.e., instantaneous) qualitative transformation of the medium from transparent to colored. Or, to put it in typically Aristotelian terms, the transparent medium serves as the material cause, and color the formal cause of vision. To explain this qualitative transformation in physical terms Aristotle falls back upon the seal-and-wax analogy offered earlier by Democritus (see passage 1.3 above). “Just as that which produces local movement [i.e., spatio-temporal motion] causes a change extending to a certain point,” Aristotle says in *On the Soul*, III, 12, 434b29-435a10 (p. 691):

[1.5] so it is also in the case of alteration, except that the agent produces it without the patient's changing its place. Thus, if an object is dipped into wax, the movement goes on until submersion has taken place, and in stone it goes no distance at all, while in water the disturbance goes far beyond the object dipped: in air the disturbance is propagated farthest of all, the air acting and being acted upon, so long as it maintains an unbroken unity. . . . On a smooth surface [such as that of the aerial medium and the eye] the air possesses unity; hence it is that it in turn sets the sight in motion, just as if the impression on the wax were transmitted as far as the wax extends.

Clearly, Aristotle does not mean to imply that the visual impression is actually stamped in the eye in a crassly mechanical way. The seal-and-wax analogy, like that of the rippling effect in water when a stone is dropped into it, is just that—an analogy, not a physical description. Another, perhaps better way of understanding the physical alteration of the transparent medium and the eye by color is by analogy to the freezing of water. Because freezing entails a change of state rather than motion or mechanical impulse, it follows “that water should be frozen simultaneously in every part” (*Sense and Sensibilia*, 6, 447a2-3 [p. 709]). Like freezing, therefore, the transmission of color occurs everywhere instantaneously.

The Extramissionist Alternative: Whereas the intromissionists cast the eye in an essentially passive role, as recipient of visual impressions, their extramissionist opponents—led by the so-called Pythagoreans—took the opposite tack, emphasizing the activity of the eye in gathering those impressions for itself. For instance, as quoted by Aristotle in *Sense and Sensibilia*, 2, 437b26-438a3 (p. 695), Empedocles (fl. c. 460 B.C.), who was not a Pythagorean, explains vision in the following terms:

[1.6] As when one who purposes going abroad prepares a lantern, a gleam of fire blazing through the stormy night, adjusting thereto, to screen it from all sorts of winds, transparent sides, which scatter the breath of the winds as they blow, while, out through them leaping, the fire, i.e. all the more subtle parts of this, shines along his threshold with incessant beams: so the primaevael fire, fenced within the membranes and delicate tissues [of the eye] give birth to a round-eyed daughter—tissues bored through with wonderful channels—and these fended off the deep surrounding flood, while letting through the fire, i.e. all its more subtle part.

Empedocles therefore likens the eye to a lantern, which, endowed with an internal fire, sheds its own light on external objects to render them visible. There is, of course, empirical evidence to support the idea that the eye has innate light or fire. The early Pythagorean, Alcmaeon (fl. c. 490 B.C.), is said

by Theophrastus to have observed that “the eye obviously has fire within, for when one is struck [this fire] flashes out” (*On the Senses*, 26, p. 89). Many ancient authorities found additional support for the theory of intraocular fire in the apparently self-luminous nature of the eyes of nocturnal animals, such as cats. And the fact that myopes squint in order to clarify their vision strongly implies that sight is due to a material efflux from the eye, as the following passage from the Pseudo-Aristotelian *Problems*, XXXI, 16, 959a3-8 (p.1510), illustrates:

[1.7] Why do the short-sighted bring their eyelids close together when they look at anything? Is it due to the weakness of their sight,³ so that, just as a man in looking at a distant object puts his hand up to his eyes, they close the eyelids to look at objects near at hand? They do so in order that the vision may proceed forth in a more concentrated form, since it passes through a narrower opening, and that it may not be immediately dispersed by passing out through a wide aperture.

The implication here is obvious: in order to see more clearly, the nearsighted must focus the stream of ocular fire more narrowly than usual in order to highlight distant objects.

Intermediate Theories: Actually, it is difficult to tie particular theorists (with the possible exception of Euclid and Hero of Alexandria) to the unmitigated extramissionist position implied in the preceding passages. Most “extramissionists,” including Empedocles, seem to have taken a *via media* between extramissionist and intromissionist extremes. Take Plato’s mythopoeic account of vision in *Timaeus*, 45b-c (pp. 152-53):

[1.8] Such fire as has the property, not of burning, but of yielding a gentle light, [the gods] contrived should become the proper body of each day. For the pure fire within us is akin to this, and they caused it to flow through the eyes, making the whole fabric of the eye-ball, and especially the central part (the pupil) smooth and close in texture, so as to let nothing pass that is of coarser stuff, but only fire of this description to filter through pure by itself. Accordingly, whenever there is daylight round about, the visual current⁴ issues forth, like to like, and coalesces with it and is formed into a single homogeneous body in a direct line with the eyes, in whatever quarter the stream issuing from within strikes upon any object it encounters outside.

Although the ocular fire-flux serves as the primary vehicle of sight in this account, it must conspire with daylight—a similar form of fire—in order to be effective. The resulting composite then washes up against external objects to render them visible. (Note that, as for Aristotle, so for Plato, light serves as a catalyst rather than an immediate cause of vision). But for Plato vision

involves more than a conspiracy between visual flux and daylight. As is evident from the following passage in *Timaeus*, 67d-e (pp. 276-77), particulate emissions from external bodies play a role as well:

[1.9] The particles that come from other bodies and enter the visual ray when they encounter it are sometimes smaller, sometimes larger than those of the visual ray itself; or they may be of the same size. Those of the same size are imperceptible—'transparent', as we call them. The larger, which contract the ray, and the smaller which dilate it, are analogous to what is cold or hot to the flesh, and again to what is astringent or burning ('pungent' as we call it) to the tongue. These are black and white, affections which are due to these particles and are similar in character, though occurring in a different field and for that reason presenting themselves in a different guise. The names should be assigned accordingly: 'white' to what dilates the visual ray, 'black' to what contracts it.

So color-perception is ultimately dependent on the tripartite physical interaction of visual flux, light, and material emanations from visible objects.

In certain ways the Stoic theory of vision is reminiscent of Plato's in that it, too, supposes a fundamental interaction between ocular flux and external light. In the case of the Stoics, however, the ocular flux (or *pneuma*) does not travel to the object but, rather, to the interface between the ocular surface and aerial medium enveloping it. When that medium is properly illuminated, the impinging visual flux transforms it immediately into a percipient extension of the eye. It is through this percipient extension that outlying objects are visually sensed, much as physical bodies are felt by means of a walking-stick. Galen offers a variant of this theory in *On the Opinions of Hippocrates and Plato*, VII, 5 (p. 455).

[1.10] At the time when we look at something, the surrounding air becomes for us the kind of instrument that the nerve in the body is at all times. It seems that the effect produced on the air around us by the emission of the *pneuma* is of the same sort as the effect produced on it by the light of the sun. For sunlight, touching the upper limit of the air, transmits its power to the whole; and the vision that is carried through the optic nerves has a substance of the nature of *pneuma*, and when it strikes the surrounding air it produces by its first impact an alteration that is transmitted to the furthest distance—the surrounding air being, of course, a continuum, so that in a moment of time the alteration spreads to the whole of it.⁵

In VII, 7 (p. 475) of the same work, though, Galen cautions against likening the percipient extension of pneumatized air to a walking-stick because:

[1.11] This . . . kind of discernment [by a walking-stick] is of resistant bodies, and it is besides more inferential than perceptive, whereas the per-

ception of our eyes is not perceptive of a thing as close packed, or of its hardness or softness, but of its color, size, and position, and none of these can be discerned by a walking-stick.

As adapted by Galen, the Stoic theory thus depends upon a fundamental alteration of the continuous aerial medium by sunlight and a concomitant alteration by the visual pneuma that strikes it at the ocular surface. With these two alterations complete, the medium provides both a physical and perceptual bridge through which the eye makes visual contact with external objects. Like Aristotle, then, the Stoics and Galen emphasized the need for a medium through which sensible alterations, not material or mechanical effects, could occur instantaneously.

The Visual Ray Theory: At a superficial level it would appear that visual-ray theorists took a radically extramissionist position in their account of vision. For instance, Euclid seems satisfied to explain vision in straightforward terms of physical contact between visual flux and external objects. Once blocked by such objects, the flux feels them visually and thereby enables the viewer to assess their spatial characteristics. Nowhere does Euclid mention color-perception or light, nor does he address the issue of how visual impressions are transmitted back to the eye for perceptual judgment. His primary concern thus seems to have been to establish the principle of rectilinear radiation on minimal physical grounds in order to set the stage for a geometrical analysis of sight.

Ptolemy, on the other hand, did address the issues ignored by Euclid. Unfortunately, the details of his physical account have been lost with the first book of the *Optics*, but there are sufficient clues scattered throughout the remainder of the work to permit us to reconstruct his explanation in a general way. For one thing, it is clear that his theory of vision was not monocausal; like Plato and the Stoics, he recognized the need for external illumination if vision was to be effective. This point is illustrated in the following passage from *Optics*, II, 4 (p. 71):

[1.12] Now luminous compactness (*lucida spissa*) is what is intrinsically visible, for objects that are subject to vision⁶ must somehow be luminous, either in and of themselves or from elsewhere, since that is essential to [the functioning of] the visual sense; visible objects must also be compact in substance in order to impede the visual flux, so that its power may enter into them rather than pass through without incident effect. Thus, it is impossible for anything to be seen without these two conditions being met, nor [can anything be seen] when one of them is met without the other.

For vision to occur, the visible object must have enough substance to be felt by

the impinging flux, provided that it is properly illuminated. But luminous compactness is not itself visible; color is the sole property of external objects accessible to sight. Thus, as Ptolemy puts it in *Optics*, II, 13 (pp. 74-75):

[1.13] A [sole] proper sensible can be found that is appropriate to each of the senses; e.g., the quality of “resisting the hand” for touch, savors for taste, sounds for hearing, and odors for smell . . . [but] color is perceived by no sense but sight. Thus, color must be the proper sensible for sight, and that is why color is taken to be what is primarily visible . . .

Color, for Ptolemy, is not a mere subjective phenomenon arising from a physical interaction between visual flux and external objects—as, for instance, Plato contends (see passage 1.9 above). Ptolemy is emphatic in according color absolute objectivity, a point that comes through clearly in the following passage from *Optics*, II, 12 and 16 (pp. 74 and 76):

[1.14] Since light and visual flux strike the surfaces of bodies together, it is quite appropriate that the first thing to be sensed in all visible objects is a characteristic of their surfaces. And color is more properly attributed to surface than to the interior of things. For this reason, the ancients used to equate surface and color, because color is a certain property affixed to the substance of an illuminated thing, and the genus “surface” is like that; and so it is an apt designation for it . . . It is therefore obvious from what we have said that color truly inheres in these objects and belongs to them by nature, and it is seen only when light and visual rays combine to make it effective. [Hence] either something that is self-luminous or something that is rendered luminous in some similar way should highlight colors by mixing with them at the surface, because [light] is generically related to none of the visible properties but color. It relates to itself, however, as [if providing] “form” to the “matter” of color.⁷

From the preceding two passages it is evident that Ptolemy follows Aristotle closely in defining color not only as the proper object of sight but also as an inherent quality of external objects. Moreover, he is aligned with Aristotle (as well as Plato and the Stoics) in reducing light to the status of mere catalyst for vision rather than a visible entity in its own right. It is, in fact, nothing more than “[the color of] an object that shines from inherent whiteness or that is exceedingly polished, for each of these is a case of brightness, and brightness is a kind of luminosity” (*Optics*, II, 5, pp. 71-72).

Whatever his apparent affinity with Aristotle, Ptolemy clearly departs from him in positing visual flux as the primary vehicle of sight. Yet, as the following passage from *Optics*, II, 23 (pp. 79-80) shows, this departure is not as radical as it may at first appear:

[1.15] According to what we have presupposed, we see any luminosity or

color by means of a passion arising in the visual flux Indeed, the passion arising in the visual flux is [called] “illumination” or “coloring.” Illumination by itself, however, is a sort of excess-condition in luminous objects, so it hurts and offends the [visual] sense. Illumination is also created along with coloring in objects that are struck by light from outside. Light and color are also transformed into one another by a transition of one into the species of the other, since luminosity provides the genus for both. And if light falls upon it, color becomes luminous, while light, if it is colored, is obviously altered. The visual flux, on the other hand, provides nothing qualitative to either of them, for it is necessary that the sense [of sight, which is] perspicuous, should have no qualification but should be pure and should suffer the qualification [passed to it] by light and color, because it shares their genus. Nevertheless, it does undergo a straightforward qualitative alteration from all colors and light.

Once it has established physical contact with an illuminated object, visual flux provides the means by which the object’s color is conveyed, through a sort of “passion,” to the eye. Being “perspicuous” and, therefore, lacking intrinsic visible qualities, the visual flux serves the very same function as the transparent medium adduced by Aristotle: it sustains the qualitative transformation that is generated by the visible color and passed to the eye. Like the Stoics, to whom he is certainly indebted, Ptolemy bases his account of vision upon both extramissionism (in the establishment by visual flux of a physical link between eye and object) and intromissionism (in the transmission of the resulting visual impression back to the eye through that link). Like Aristotle, however, he grounds that visual impression in color, and color alone.

Color and Color-Perception: To this point we have looked at two apparently contradictory theories of color in antiquity: the Platonic theory, according to which color seems to be absolutely subjective, a psychological effect created by dilation and contraction of the visual flux (see passage 1.9 above), and the Aristotelian theory, according to which color seems to be absolutely objective, an inherent quality of things that is independent of any psychological agency (see passage 1.4 above). Yet, despite these real—and fundamental—differences, the two theories were based on the same crucial assumption, one that was in fact common to virtually every ancient thinker: namely, that black and white are the polar constituents of all colors. Thus, for the most part, colors are differentiated according the relative proportion of black and white that go into their formation; the whiter, the more vivid the color, and, by extension, the blacker, the less vivid it is. The following passage from Aristotle’s *Sense and Sensibilia*, 3, 439b20–440a12 (p. 698) illustrates this supposition quite clearly:

[1.16] It is conceivable that the white and the black should be juxtaposed in quantities so minute that either separately would be invisible, though the joint product would be visible; and that they should thus have the other colours for resultants. Their product . . . must be of a mixed character—in fact, a species of colour different from either. Such, then, is a possible way of conceiving the existence of a plurality of colours besides the white and black; and we may suppose that many are the result of a ratio.

Another [way to explain the genesis of intermediate colours] is that the black and the white appear the one through the medium of the other, giving an effect like that sometimes produced by painters overlaying a less vivid upon a more vivid colour, as when they desire to represent an object appearing under water or enveloped in a haze, and like that produced by the sun, which in itself appears white, but takes a crimson hue when beheld through a fog or a cloud of smoke.

But what a given color is can be quite different from what it appears to be. For instance, colors can be perceived quite differently according to the light under which they are seen. Likewise, they can appear different when seen through such media as water or mist. Thus, as the matter is put in the Pseudo-Aristotelian *On Colours*, 3, 793b14-21 (p. 1222):

[1.17] We never see a colour in absolute purity: it is always blent, if not with another colour, then with rays of light or with shadows, and so it assumes a tint other than its own. That is why objects assume different tints when seen in shade and in light and sunshine, and according as the rays of light are strong or weak. . . . Again they vary when seen by firelight or moonlight or torchlight, because the colours of those lights differ somewhat. . . . Thus all hues represent a threefold mixture of light, a translucent medium (e.g., water or air), and underlying colours from which the light is reflected.

One more variable must be taken into account if we are to evaluate color-perception in all its complexity. Just as the quality of illumination from external light-sources, such as the sun, the moon, or fires, can modify color-perception, so can the quality of visual illumination. And the primary determinant of that quality is the strength or weakness of the visual radiation, which in turn varies with a number of factors, distance being the most salient. Accordingly, the weaker the visual illumination, the less vivid the resulting color-perception; but since vividness determines hue, then a given color seen under these conditions will appear of a different hue than it should. As we will see in our study of the rainbow in Section V below (see esp. passage 5.2), this model of color-perception and its modification by the weakening of visual radiation provides the key to understanding Aristotle's analysis of the color-bands in the rainbow.

All in all, then, even the most objective accounts of color in antiquity—Aristotle's being the most salient example—were moderated by a clear understanding that color-perception depends upon a variety of factors beyond the mere quiddity of the color seen. The quality of light under which the color is seen, the distance of the observer from the color, the medium through which the color is viewed, all played a part in how the given color would be apprehended by the perceiver.

2. *The Physiology of Vision*

Among the Greek optical theorists currently known to us, no one before Galen offered a systematic and coherent functional account of the eye as a visual instrument. To be sure, pre-Galenic theorists referred in passing to basic features of the eye and its internal structure, but it is impossible on the basis of such occasional references to get a clear picture of how these theorists understood the eye and its specific function in sight. Empedocles, for instance, is reported by Theophrastus, in *On the Senses*, 7 (pp. 71-73), to have held that:

[1.18] perception occurs because something fits into the passages of the particular [sense organ]. For this reason the senses cannot discern one another's objects, he holds, because the passages of some [of the sense organs] are too wide for the object, and those of others are too narrow. . . . Then he attempts to tell us the character of the organ of vision. Its interior, he says, is of fire; while round about this [internal fire] are earth and air, through which the fire, by reason of its subtilty, passes like the light in lanterns. The passages [of the eye] are arranged alternately of fire and of water: by the passages of fire we perceive white objects; by those of water, things black; for in each of these cases [the objects] fit into the given [passages].

Democritus, on the other hand, emphasizes the watery nature of the eye, for it is by virtue of its moistness that the "reflection," or *emphasis*, that produces vision is effected at the corneal surface (see passage 1.3 above). The visual percipience of the eye would thus seem to be tied to the intrinsic moistness that makes it receptive to the visual impressions created by *emphasis*. Furthermore, if *emphasis* is the primary mechanism of vision, the cornea must represent the point of initial visual sentience.

Aristotle agrees with Democritus that the watery nature of the eye gives it its peculiar receptivity to visual impressions but disagrees with him about the mechanism of impression. As he puts it in *Sense and Sensibilia*, 438a5-16 (pp. 695-96):

[1.19] Democritus . . . is right in his opinion that the eye is water; not, however, when he goes on to explain seeing as mirroring . . . True, then, the

visual organ is composed of water, yet vision appertains to it not because it is water, but because it is transparent—a property common alike to water and to air. But water is more easily confined and more easily condensed than air; [so] the pupil, i.e., the eye proper, consists of water.

So far, fire and water have been cited as the instrumental constituents of the eye. Diogenes of Apollonia (fl. c. 430 B.C.) offers yet a third possibility, which is described by Theophrastus in *On the Senses*, 39-40 (p. 101):

[1.20] Diogenes connects the senses with the air, even as he connects it with both life and thought. He would accordingly seem to ascribe [perception] to likeness; for, he holds, there would be neither activity nor passivity unless all things were from a single [source]. Smelling is effected by the air about the brain; since the air is massed there and is commensurate with odour; while the brain of itself, with its ducts, is already of light consistency Sight arises when objects are reflected in the pupil, but it occasions perception only when mingled with the internal air. This is capable of proof: for if the ducts [the optic nerves?] become inflamed, there is no union with the internal [air], and sight is impossible, although the image is still there as before.

Two aspects of this account are worth noting. First, unlike the other theorists cited, Diogenes links the eye itself to the seat of perception (in the brain), using air as the crucial mediating entity. In short, he offers a basic, although vague, anatomical model of visual perception. Second, Diogenes emphasizes the distinction between vision, as a mere physical process (through *emphasis*), and vision as a perceptual process (through mingling of the emphatic impression with the internal air that connects the eye and brain).

Suffice it to say, the accounts given to this point offer no systematic physiological model of vision. For that we must leap several centuries ahead to Galen, who deals with the anatomy and physiology of vision at great length in *On the Usefulness of the Parts*. Let us start with his description of the brain in VIII, 6 (p. 398) of that work:

[1.21] In substance the encephalon [i.e., the brain] is very like the nerves, of which it was meant to be the source, except that it is softer, and this was proper for a part that was to receive all sensations, form all images, and apprehend all ideas. For a substance easily altered is most suitable for such actions and affections.

The brain itself has an internal structure, consisting of various hollows, or ventricles, that are connected to one another throughout. It is within and through these ventricles that psychic pneuma, the “material cause” that underlies the act of reasoning, is formed and diffused. Thus, as Galen puts it in *On the Usefulness of the Parts*, IX, 4 (p. 432):

[1.22] [Elsewhere] I have given the demonstrations proving that the rational soul is lodged in the encephalon; that this is the part with which we reason; that a very large quantity of psychic pneuma is contained in it; and that this pneuma acquires its own special quality from elaboration in the encephalon.

Thus, although the brain is the seat of reason, as well as of sensation, it is psychic pneuma that is the ultimate medium for sensitive and intellectual functions (note the affinity with Diogenes' account in passage 1.20 above). Psychic pneuma, in its turn, is generated from vital pneuma that is carried to the brain through the carotid arteries to the retiform plexus at the base of the brain.⁸ "Here," concludes Galen in *On the Usefulness of the Parts*, IX, 4 (pp. 432-33):

[1.23] we see that both the retiform plexus and other features of its construction are in wonderful harmony For the whole encephalon is interwoven with these intricately divided arteries, many of whose branches end at its ventricles; [and] remaining for a very long time in the arteries, the [vital] pneuma is elaborated, but when its elaboration is complete, it falls at once into the ventricles of the encephalon; for it ought not to be delayed longer, nor should it escape before it is elaborated.

Much like the heart, with its venous and arterial connections, the brain is linked to the body through a complex network of nerves. Part of this network consists of the two optic nerves, which are hollow throughout. Branching off from the front of the brain, these two conduits join at the optic chiasma and, after being re-separated, reach their respective eyes, to which they are attached at the rear. The ocular globe itself consists of various humors, among which the vitreous humor fills most of the globe. This humor is adjoined toward the front of the eye by the crystalline humor, which is encased within the lens. Between the lens and cornea, finally, lies the aqueous humor. The outer shell of the eye consists of various tunics, from the sclera and cornea on the very outside, to the choroid, iris (whose perforation at the center forms the pupil), and retiform tunic (i.e., retina) inside. These tunics serve a variety of functions from protection to nutrition. With these anatomical features in mind, let us look at Galen's summary of visual physiology in *On the Usefulness of the Parts*, IX, 6 (pp. 402-403):

[1.24] It is absolutely necessary for each of the [sense organs] to be altered if sensation is to occur. They are not, however, all altered by every perceptible thing; rather, the bright luminous sense instrument [i.e., the eye] is altered by colors, the airlike instrument [i.e., the ear] by sounds, the vaporous instrument [i.e., smell] by odors, and in a word, like is perceptible by like. The airlike sense instrument can never be altered by colors, for if an instrument

is to undergo alteration by colors easily and simply, it must be radiant pure and bright Hence, none of the sense instruments except the instrument of vision will be altered by colors, for vision alone has a sense instrument that is radiant, pure, and glistening, namely, the crystalline humor But it would be of no use for this alteration to take place unless it was recognized by the ruling principle which forms images, remembers, and reasons. Accordingly, the encephalon extends a part of itself [via the optic nerves] to the crystalline humor in order to know how it is being affected, and this outgrowth is properly the only one to have a perceptible channel, because it alone contains a very large amount of the psychic pneuma Surely, then, . . . since the sense instrument of vision must be made bright and radiant, a great deal of pneuma is properly sent to it from the source, and from the encephalon itself there is extended to it a pure, unadulterated offshoot [i.e., the system of optic nerves]. . . . As regards the eyes, even though they are covered as closely as possible on all sides, alteration from external colors easily reaches the part of the encephalon contained in them. For the [cornea] is so thin, clear, and pure, that it closes off neither this part nor the alteration that passes through it, and after the cornea, directly at the pupil, comes the crystalline humor, to which the part of the encephalon contained in the eye is adnate.

The “ruling principle” to which Galen adverts in the preceding passage is essentially the same as the Stoic *hegemonikon*, which is the psychological faculty that directs and assesses all animal and intellectual functions. Using psychic pneuma as its medium of activity, this ruling principle is ultimately responsible for making sense of the visual information passed to it through the optic pathway from the crystalline lens. The anterior surface of the crystalline lens has thus replaced the cornea as the locus of initial visual sentience. That sentience, in turn, is animated by the psychic pneuma that is diffused throughout the optic system from the “source” in the brain.

The physiological model of vision that Ptolemy seems to have had in mind is consistent, at least in a general way, with that of Galen. Thus, although he apparently accepts the cornea rather than the crystalline lens as the seat of initial visual sentience, he adduces a Governing Faculty (*virtus regitiva*) to control visual perception through a “nervous principle.” There are strong indications, moreover, that he thinks in terms of a physiological pathway for sight. Accordingly, he lodges the immediate source (*principium*) of visual flux at the center of the eye but put both eyes under control of an ulterior source (*principium* = Apex) that may lie in the optic chiasma. Ptolemy therefore seems to have understood the visual pathway according to a hierarchy of agent-sources, from the Governing Faculty in the brain, through the controlling source at the optic chiasma, to the immediate source of visual flux at the center of the eye.⁹ Accordingly, he must have conceived of the visual process

in terms of stages along that pneumatically-charged pathway. If not Galenic in actual inspiration, Ptolemy's model of visual physiology is certainly Galenic (perhaps "Alexandrian" is more apt) in spirit.

3. *The Psychology of Visual Perception*

The Aristotelian Account: As with the physiology of vision, so with the psychology of visual perception, few of the extant sources offer a systematic account of how visual information is assimilated and assessed by the sensitive and rational faculties. Aristotle's *On the Soul* is one of those few. In it he offers a clear schema, according to which vision is analyzed in three basic levels: brute sensation, perception, and apperception. Let us start with an analogy offered by Aristotle in *On the Soul*, II, 1, 412b18-22 (p. 657):

[1.25] Suppose that the eye were an animal—sight would have been its soul [i.e., its animating principle], for sight is the substance or essence of the eye which corresponds to [its functional nature], the eye being merely the matter of seeing; when seeing is removed the eye is no longer an eye, except in name—it is no more a real eye than the eye of a statue or of a painted figure.

The eye can thus be considered in two ways. On the one hand, as a mere physical entity, it is nothing more than its constituent matter. On the other hand, as a sentient entity, it is endowed with a certain level of animacy that permits it to "feel" the color-impression that arouses sight. Moreover, Aristotle continues in *On the Soul*, II, 5, 417a12-418a6 (pp. 663-65):

[1.26] 'sense' . . . must have two meanings, sense potential and sense actual. Similarly, 'to be a sentient' means either (a) to have a certain power or (b) to manifest a certain activity. . . . [Therefore] what has the power of sensation is potentially like what the perceived object is actually; that is, while at the beginning of the process of its being acted upon the two interacting factors are dissimilar, at the end the one acted upon is assimilated to the other and is identical in quality with it.

As a sentient entity, then, the eye has the potential to grasp its proper sensible, color, in such a way that, in actualizing that potential, it "becomes" the sensible and, in so doing, also actualizes the potential of its proper sensible to be seen. Brute visual sensation is therefore limited to the perceptual grasp of color in a fundamentally non-judgmental way. "Each sense," Aristotle concludes in *On the Soul*, II, 6, 418a14-17 (p. 665), "has one kind of object which it discerns, and never errs in reporting that what is before it is colour or sound (though it may err as to what it is that is coloured or where that is, or what it is that is sounding or where that is)." Properly speaking, then, the eye sees color only in a generic sense; it makes no judgment about whether the

color is red or green.

If visual perception were limited to color alone, of course, our apprehension of external reality would be far more impoverished than it is. So how is it that we “see” so many other physical characteristics of that reality? In response to this question, Aristotle adduces two additional types of sensibles: common and incidental. The common sensibles, he informs us in *On the Soul*, II, 6, 418a17-25 (p. 665):.

[1.27] are movement, rest, number, figure, magnitude; these are not peculiar to any one sense but are common to all. [On the other hand] we speak of an incidental object of sense where e.g. the white object which we see is the son of Diaries; here because ‘being the son of Diaries’ is incidental to the directly visible white patch we speak of the son of Diaries as being (incidentally) perceived or seen by us. Because this is only incidentally an object of sense, it in no way as such affects the senses. Of the two former kinds [i.e., special and common sensibles], the first kind—that of special objects of the several senses—constitute *the* objects of sense in the strictest sense of the term and it is to them that in the nature of things the structure of each several sense is adapted.

The common sensibles are therefore not actually sensed; they are implicit in (and thus conveyed by) the special sensibles and inferred mediately by a faculty of the soul that Aristotle calls the “common sensibility” (*aisthesis koine*). This faculty, which possesses no specific organ, gathers the special sensibles together and, from them, forms a perceptible composite or “image” for perceptual scrutiny. On the basis of such images, finally, we judge the represented objects to be “the son of Diaries,” “Socrates,” “the sun,” or whatever. In other words, just as the special sensibles convey the common sensibles, so the common sensibles convey the incidental sensibles. It is at this level of perceptual scrutiny, of course, that error can arise in the form of misjudgments about what is being represented.

To elucidate Aristotle’s three-stage account of visual perception, let us take as an example the conclusion that “the white object which we see is the son of Diaries.” Suppose, first, that the eye actualizes the potential visibility of the color-impression passed to it from Diaries’ son. Thus “seen,” the resulting image can be taken in two ways. At a physical level, it is nothing more than a color-presentation. At a visual level, however, it is a depiction, so that, like a painting, it conveys a variety of spatial characteristics that are not actually present in it. Still, all we actually see at this level is the color-presentation itself, nothing else. Or, to return to the example at hand, the resulting visual impression consists of “white object.” It is as a depiction, however, that this impression is passed to the common sensibility, which realizes the ulterior spatial characteristics (i.e., the common sensibles) implicit in the visual im-

pression. Accordingly, the common sensibility creates from the mere visual representation a more elaborate internal image consisting of, say, "white object of such-and-such a relative size and shape." Like the visible representation from which it is created, this perceptible representation conveys certain characteristics that are not actually present in it. Such characteristics (the so-called incidental sensibles) are eventually elicited through mental scrutiny, the result being an intellectual judgment that the "white object of such-and-such relative size and shape" before us is in fact "the son of Diareos."

Ptolemy's Account: Although Ptolemy's account of visual perception is consonant with Aristotle's, Ptolemy is far more concerned than Aristotle with the problem of visual illusions arising from perceptual misjudgments. Indeed, it would be fair to say that Ptolemy's ulterior purpose in the *Optics* as a whole is to correct such misjudgments by explaining them away. Take, for instance, the case of reflection in a plane mirror. The very fact that we "see" the object (or, more precisely, its image) behind the mirror represents a visual illusion based on perceptual misjudgment. To correct this misjudgment, we must know precisely how mirror-images are formed in relation to ourselves, as viewers, the mirror, and the represented object. As we will see in due course, the fundamental corrective is to be found in the law of equal angles and the principle of image-location.

Bearing this point in mind, let us start with Ptolemy's discussion of the visible properties in *Optics*, II, 2-3 (p. 71):

[1.28] The visual faculty apprehends corporeity, size, color, shape, place, activity, and rest. Yet it apprehends none of these without some illumination and something [opaque] to block the passage [of the visual flux]. We need say no more on this score but must instead specify what characterizes each of the visible properties We contend, therefore, that these visible properties exist in two ways, one of which depends upon the disposition of the visible property [itself] and the other upon the action of the visual faculty.

Perhaps the most salient feature of Ptolemy's list of visible properties is that it includes four of Aristotle's "common sensibles,"—size (or magnitude), shape (or figure), activity (or movement), and rest—all of which, with Ptolemy's addition of place, define spatial characteristics. Unlike Aristotle, however, Ptolemy makes a clear distinction between the objective grounds ("the disposition of the visible property") and the subjective grounds ("the action of the visual faculty") of sight.

Taken objectively, Ptolemy continues in *Optics*, II, 3-6 (pp. 71-72), the visible properties can be distinguished in three ways:

[1.29] some . . . are intrinsically visible, some primarily visible, and some secondarily visible. Now luminous compactness is what is intrinsically visible, for objects that are subject to vision must somehow be luminous, either in and of themselves or from elsewhere, since that is essential to [the functioning of] the visual sense; visible objects must also be compact in substance in order to impede the visual flux, so that its power may enter into them rather than pass through without incident effect On the other hand, colors are primarily visible, because nothing, besides light, that does not have color is seen. Still, colors are not intrinsically visible, since colors are somehow contingent on the compactness of bodies and are not visible *per se* without light All the rest of the aforementioned visible properties are secondarily visible, because the visual faculty apprehends things as bodies by means of their [inherent] colors and characteristics, whereas objects that have no compactness, but are exceedingly tenuous and have no color are neither sensed nor perceived as bodies by the visual faculty. Furthermore, size, place, and shape are perceived only through the mediation of bodies' surfaces, which coincide with the colors upon which external light falls. Activity and rest, as well, are apprehended by means of an alteration, or lack thereof, in any of the aforementioned visible properties.

Luminous compactness is therefore a precondition for seeing. Color, and color alone, is the proper sensible for sight, the remaining "sensibles" being conveyed mediately by it. As if to emphasize this mediating role, Ptolemy continues his discussion of color in *Optics*, II, 7, (p. 72):

[1.30] While the faculty of sight apprehends illuminated colors immediately, it apprehends the rest of the visible properties by means of such illuminated colors, not insofar as they are colors but only insofar as they have boundaries. For the visual faculty apprehends shapes and dimensions by means of the boundaries of the colored object, while place is apprehended by means of its location. The visual faculty also apprehends the motion or rest of these same colors by means of their change or lack thereof. And the motion or rest of shapes, dimensions, and location is perceived by means of the motion or rest of the boundaries or places of the colored object.¹⁰ For instance, an object that appears white certainly does not appear round, small, near, or stationary on account of its whiteness, just as the characteristics of a resisting object that the hand feels are not grasped through its hardness. On the contrary, these properties are discerned by means of either the boundary of whiteness or the boundary of hardness—the boundary in this case being an essential characteristic [of that hardness] or its size according to its extended nature.

Given how we perceive visible objects and what it is about them that we perceive, it follows that, if visual perception is to be veridical, certain threshold conditions must be met. Otherwise, the object, or certain of its defining

characteristics, will escape sight. For instance, if the boundaries delimiting objects are inadequately clear—that is, if the colors of two abutting objects are too similar to be distinguished—the separation between the objects will not be discerned. But the primary threshold conditions have to do with visual clarity, which can vary according to a number of factors. Ptolemy discusses several of these in *Optics*, II, 18-19 (pp. 76-77):

[1.31] an object is seen more clearly when more visual flux impinges upon it or when more light shines upon it—e.g., what is looked at directly is seen more clearly than what is looked at by means of reflection or refraction. Even so, an aggregation of [uninterrupted] visual rays is weakened when they are extended out to a great distance. Also, what is looked at with both eyes is seen more clearly than what is looked at with either eye alone. And what is self-luminous is seen more clearly than what is illuminated by something else. Also, the larger or more numerous the luminous sources, the more clearly the object upon which their light shines is seen. On the other hand, among objects whose appearance depends upon the quality of [radiative] effects, those that lie directly in front of, and at right angles to the rays are seen more clearly than those that do not.¹¹ For everything that falls orthogonally strikes its subjects more intensely than whatever falls obliquely. Also, what is polished is seen more clearly than what is rough, because there is disorder in a rough object resulting from the fact that its parts are not arranged in a regular way. But the parts of a polished object have a certain regularity, and [so] brightness is inherent to it. Dense objects, as well, are more clearly seen than rare ones, because rare bodies give way to the impinging [ray], whereas dense bodies resist it. Objects that radiate [light] by themselves are also seen more clearly than those that are lit by something else, and so, for example, is an object that is seen in illuminated, rarefied air. Furthermore, objects that lie a moderate distance from the viewpoint appear more clearly [than those that do not, whereas] objects that lie far away [from the eye] appear less clearly, since the visual rays, as they stream outward, take on some of the blackness of the air through which they pass. Thus, distant objects appear nebulous, as if seen through a veil.

Variations in the flow of visual flux also play a part in visual discernment. As Ptolemy remarks in *Optics*, II, 86 (p. 107):

[1.32] It is because of an abundance of visual power that objects are seen at a distance. Hence, older people always look at an object up close, because, along with the rest of their faculties, the visual power is produced more weakly in them. On the other hand, those who have deep-set eyes see farther than those who do not have such eyes; and the reason for this is that their visual power is compressed, for when it emanates from narrow places, the visual flux is stretched and elongated.¹²

Proper visual perception requires adequate time as well. Take the example

given by Ptolemy in *Optics*, II, 96 (p. 110), of a spinning potter's wheel that is daubed with various colors:

[1.33] In falling on all the [constituent] colors, the [visual flux] cannot distinguish between first and last nor among those that [now] occupy different locations. In fact, all the colors appear simultaneously throughout the disc as a single, uniform color that is the same as the color that would actually be formed from a mixture of the constituent colors. By the same token, if spots of a color different from that of the disc are marked on it (provided they are not on its very axis), they will appear to form circles of the same color [as the given spot] when the disc is rapidly spun.

All in all, then, clarity of visual perception is affected not only by physical factors, such as too little (or too much) illumination or a weak flow of visual spirit, but also by temporal and spatial factors, such as too little or too much distance between eye and object,¹³ or too little time to perceive correctly. In all these cases, though, the cause of misperception is external, not internal. In short, the visual faculty itself is not to blame.

As to misperceptions or illusions that are due to the visual faculty itself, Ptolemy divides them into three basic types. The first type involves illusions that are caused by physical alterations in the visual flux. Among these illusions, the most striking is the so-called oculogyral illusion, which is discussed in *Optics*, II, 121 (p. 119):

[1.34] A particular kind of continuous sweep of the visual [flux] gives the impression that the visible object is moving. This sort of sweep sometimes occurs at the very source of the visual flux, as happens, for instance, in the case of vertigo and fainting, whose effect reaches to the eye itself. For, while striking the visible object, the ray continues to follow the motion of its source, which has changed the direction of its focus, but the sense does not detect this sort of [internal] motion.

Assume that a viewer spins clockwise for awhile, then stops. Evidently, what happens in this case is that, once the original motion ceases, the flux at the source (presumably at the center of the eye) continues to spin inertially in a clockwise direction. Transmitted to the flux emanating from the eye, this inertial spin causes the flux to pass over those objects in a clockwise direction. Knowing himself to be at rest, yet not sensing the inertial motion of the flux, the viewer interprets the clockwise scan of the flux in terms of a counterclockwise motion among the surrounding objects.

Reflection and refraction also distort the visual flux, causing a misperception of the object's location. Thus, as Ptolemy observes in *Optics*, II, 115 (pp. 117-118):

[1.35] In the case of reflection and refraction, . . . the object appears [to lie]

directly in line with the [incident] visual ray, even though this is not actually the case insofar as the ray is broken. Moreover, among objects that are seen in this way, some appear to lie closer than they actually are—the true measure being the distance between the objects and the viewers in combination with the distance between the objects and the surface at which the breaking occurs. Other objects, however, appear to lie at the same distance as the true one, and others yet appear to lie at a greater distance [than the true one].¹⁴ In addition, certain objects are seen by those visual rays that lie on the same side with them, whereas others are seen by different rays, depending on the surface-shapes of the bodies that cause the breaking; this is the case, for example, when we see left-hand objects with right-hand rays or higher objects with lower rays, and vice-versa.¹⁵

Since the breaking of the ray is not felt, the image is seen along the continuation of the incident ray rather than where it actually is. Moreover, the shape of the reflecting surface can distort the image that appears along the incident ray in various ways, so that the image misrepresents the object that generates it. The resulting misapprehension of the object is therefore due to the way in which the visual flux is reflected from the given surface.

The second basic type of illusion arises from perceptual misjudgments of visual data that are presented correctly by the visual flux. As an example, Ptolemy offers the following in *Optics*, II, 127-128 (p. 122):

[1.36] According to the colors applied to them, surfaces sometimes appear convex and sometimes concave. Thus, a painter who wishes to represent these two shapes by means of colors paints the part he wants to appear higher a bright color, whereas the part he wants to appear concave he paints with a weaker and darker color. This is why we judge a concave veil to be convex when we view it from afar. The reason is not that the wind disposes it in such a way that sunlight and the visual flux reach the area of concavity [blown inward by the wind]. Rather, the reason is that the [relatively] orthogonal rays strike the middle of the veil so that it shows forth vividly, whereas at its outer edges either no ray at all or a somewhat oblique one strikes it, which is why it appears dark [toward the edge]. Accordingly, then, the edges of the veil appear depressed while the middle appears elevated, and this is how something that is actually convex appears.

Failure to distinguish relative motion and rest also belongs to this second type of illusion. This failure is discussed in *Optics*, II, 131 (pp. 123-124):

[1.37] When a boat stands still in a calm, waveless river that flows swiftly, anyone in the boat who does not look at the shoreline beyond [but focuses on the river] judges that the boat is moving swiftly upriver while the water is standing still. The reason for this illusion is that the motion of the water sensed by the visual flux, being opposite to that with which the boat is as-

sumed to move, is manifested by the contrast between the color of the boat and the color of the water. Now the contrast created by the motion of the parts of the water's surface alone is not clear to the sense because of the uniformity of the parts of the [water's] surface and the similarity [throughout] of its color. Yet, according to the motion of the visual flux upon the parts of the visible object's surface, it is necessary that either the water or the boat appear to move. Thus, since the water will appear calm, the motion must appear to belong to the boat. On the other hand, if we look at the water, the shoreline, and the boat all at the same time, and if we take cognizance of the fact that the shoreline is stationary, then we will see that the boat is stationary, since the boat is seen by the same rays that see the shoreline. We will also see the water moving since we will have realized that the boat and the shoreline are stationary.

The problem here is that, by focusing too closely upon a single aspect of the given phenomenon (i.e., concentrating on the water flowing past the stationary boat without looking at the shoreline), the perceptual faculty misinterprets the phenomenon as a whole (i.e., judging that the boat, rather than the water, is moving).

The third, and final type of visual illusion discussed by Ptolemy is due to false mental inference. This type of illusion is at play, for instance, in the interpretation of mirror-images. Thus, as Ptolemy explains it in *Optics*, II, 138 (p. 126), when we regard ourselves face-on in a plane mirror:

[1.38] our sight shows us our [right-hand and left-hand] sides in the way that is natural for it to show objects viewed directly: i.e., what is seen by right-hand rays appears to the right, while what is seen by left-hand rays appears to the left. Our mind, however, shows us right as left and left as right, because objects that actually face us are so disposed that their right is opposite to our left, while their left is opposite to our right. And this is why, when we move one of our hands [in front of a mirror] our sight tells us that the hand that moves [in the mirror] is the one facing it [i.e., right to right or left to left], while our mind tells us the opposite.

What differentiates this type of illusion from the second sort involving perceptual misjudgments arising from properly presented visual information is that, in this case, the information is improperly presented. That is, the image we see face-on in a mirror is presented to us as if it were of someone else facing us directly. Our *perceptual* judgment of that image is therefore based on that presentation, so that, in order to rectify the illusion, we must unpack it intellectually.

Ptolemy's analysis of visual illusions reflects his understanding of visual perception as a three-stage process running from sensation, through perception, to apperception. The first type of illusion, which occurs at the level of

brute sensation, involves “feeling” rather than judgment. This characterization is evident, for instance, in the oculogyral illusion, the result of which is a “feeling” that, even when the viewer stops spinning, his surroundings continue to rotate for awhile. Likewise, in the case of reflection, the object appears to be where it is not (i.e., directly along the line of sight) because the flux fails to “feel” the impedance of the mirror-surface that breaks it. The second type of illusion, for its part, involves the common sensibles and, therefore, the perceptible characteristics of things. The example of illusionism in painting illustrates this type of misperception perfectly. The very fact that we are able to discern spatial characteristics in two-dimensional representations (e.g., *trompe l’oeil* paintings) is a clear indication that it is not our eyes, but our brains that are deceiving us. Finally, as to the third type of illusion, involving false mental inference, it is evident that we are dealing not with common sensibles but with incidental sensibles. Thus, in the case of image-reversal, the representation that we initially “see” is not of ourselves but of someone else facing us as if through a window. The crucial feature of Ptolemy’s analysis of visual illusions is therefore not to be found in detail or specification; indeed, many of the examples cited by him were part of a common stock gathered together long before he wrote.¹⁶ It is, rather, to be found in the systematic way he fit that analysis into the broader framework of his theory of visual perception.

In his *Theories of Vision from Al-Kindi to Kepler*, the distinguished historian of science, David Lindberg, classifies Greek optical thought according to three broad (albeit sometimes overlapping) traditions. The “medical” tradition, for its part, was “concerned primarily with the anatomy and physiology of the eye and the treatment of eye disease.” The “philosophical” tradition, on the other hand, was “devoted to questions of epistemology, psychology, and physical causation.” The “mathematical” tradition, finally, was “directed principally toward a geometrical explanation of the perception of space.”¹⁷ While this schema is far from ill-conceived as a heuristic, it is nonetheless potentially misleading. For it can be taken to imply that, in limiting themselves to the analysis of spatial perception, mathematical theorists were unconcerned, or only marginally concerned, with physical, physiological, and psychological issues. Their approach to optics, in short, was instrumentalist to the core. That this is not the case, at least for Ptolemy, should be abundantly clear by now. Far from ignoring physical, physiological, and psychological issues, in fact, Ptolemy confronted them head-on in order to provide a fully-articulated, coherent theoretical context for his mathematical analysis of sight. Without a proper appreciation of this theoretical context, our understanding of the de-

velopment of mathematical optics, as outlined in the following three sections, will be not only incomplete but badly distorted.

NOTES

¹Diogenes Laertius, *Lives of the Eminent Philosophers* 10.48-49, trans. R. D. Hicks, Loeb Classical Library (Cambridge, MA: Harvard University Press, 1958-59), vol 2, pp 577-79.

²By “pupil” (Greek = *kore*) is meant the corneal surface above the pupil and through which the pupil can be seen. In fact, such a conflation of pupil with cornea was common among ancient authorities (see, e.g., passages 1.19 and 1.20 below).

³The Greek term *opsis*, which is translated by “sight” here and by “vision” a couple of lines below, has a range of meanings, including “visual flux” and “visual ray.” In Euclid’s *Optics*, for instance, *opsis* is clearly meant to be taken as “visual ray.”

⁴Here, too, the Greek term is *opsis*.

⁵Note the obvious affinity with Aristotle’s description of alteration in passage 1.5 above.

⁶The Latin term *visus*, which is translated by “vision” here, has much the same range of meanings as its Greek counterpart *opsis*. Thus, at numerous places throughout Ptolemy’s *Optics* it should be taken to mean “visual flux” or “visual ray” rather than “sight” or “eye.”

⁷As “form” to color’s “matter,” then, light is necessary for the actualization of color’s visibility. Or, to put it another way, color without light, is potentially, but not actually, visible, and so is light without color.

⁸This “retiform plexus” (the so-called *rete mirabile*) consists of an extraordinarily fine network of arteries that is most noticeable in ungulates but does not, in fact, occur in the human brain, where its counterpart is to be found in the Circle of Willis.

⁹Since most of the clues about Ptolemy’s conception of the physiology of vision are scattered throughout the *Optics*, particularly in book 3, we will simply signal them as they crop up in various passages over the course of this study.

¹⁰The perception of corporeity, which is not spatially determined, results from the very act of seeing: since compactness is a precondition of sight, the fulfillment of that precondition indicates to the viewer that what is seen is a body of some sort (cf. passages 1.28 and 1.29 above).

¹¹In other words, objects that are seen toward the outer edges of the visual field appear less distinct—and are thus perceived less clearly—than those nearer the center of that field. By “radiative effects” Ptolemy means to implicate the dynamic properties of the visual radiation which can be likened to physical projection.

¹²Note the similarity of this account to the Pseudo-Aristotelian account in passage 1.7 above of why myopes squint in order to improve their distance-vision.

¹³A standard classical example of how distance can affect visual perception is the apparent rounding of square objects when they are seen from afar.

¹⁴For instance, in convex mirrors objects appear to lie farther away than they should.

¹⁵Here Ptolemy is adverting to the image-inversion that can occur in concave mir-

rors.

¹⁶Lucretius, for example, provides a long list in *De rerum natura*, 4, 324-447.

¹⁷*Theories of Vision* (Chicago: University of Chicago Press, 1976), p. 1.

SECTION II

OPTICS PROPER: ANALYSIS OF DIRECT VISION

Well before the time of Euclid it was recognized that what we see depends upon whether the object is viewed in unmediated fashion or whether it is viewed by means of a reflecting or refracting surface. According to ray-analysis, then, the domain of optics, taken at its most general level, can be divided into three subdomains. On the one hand, optics proper is concerned with vision that occurs by means of unbroken rays. On the other, catoptrics deals with vision that occurs by means of completely broken rays. And, finally, dioptrics covers vision that occurs by means of partially broken rays. In this section we will deal with the first subdomain, starting with an account of the geometrical and perceptual properties of the visual cone. We will then move on to the perception of physical space, examining how the “common sensibles” are apprehended according to visual ray analysis. We will conclude with a discussion of binocular vision and the perceptual problems that arise from improper accommodation.

1. The Visual Cone

As we have already noted, ancient mathematical optics was grounded in the assumption that the eye radiates visual flux outward toward visible objects. The basic model of visual radiation is described succinctly in the first three definitions (actually postulates) of Euclid's *Optics* (p. 2):

[2.1] Let it be supposed: (1) that straight lines diverge outward from the eyes to comprehend vast spaces, (2) that these visual rays form a cone whose vertex is located in the eye and whose base is formed at the boundaries of visible objects, (3) that objects with which the visual flux makes contact are seen, whereas those objects not contacted by the visual flux are not seen.

In the tenth book of *On the Usefulness of the Parts*, Galen offers a more elaborate description (X, 12, pp. 492-93):

[2.2] Let there be a circle; . . . from the mid-point of the circle (which is also called its center) think of a straight path to the pupil¹ of the eye that is seeing it, a path not bending in any direction or deviating from its straight course;

think rather of that straight line as you would of a thin hair or the filament of a cobweb accurately stretched from the pupil to the center of the circle. Again, from the pupil to the line which bounds the circle and is also called its circumference imagine a series of very many other lines extending like thin cobwebs. Call the figure bounded by all these straight lines and by the circle a cone, and think of the pupil as its apex and the circle as its base.²

As propositions 2 and 3 of Euclid's *Optics* (pp. 4-6) illustrate, this model of discrete radiation offers a simple explanation for variations in visual acuity with distance:

[2.3] [PROPOSITION 2] *Of equal magnitudes lying at a distance, those lying nearer are seen more distinctly.*

Let B [in figure 2.1] be the eye and GD and KL the visible [magnitudes]. It must be borne in mind that they are equal and parallel. Let GD be the nearer, and let visual rays³ BG, BD, BK, and BL fall [upon their respective visible magnitudes]. Now we would not say that the visual rays passing from the eye to KL reach it through points G and D, for straight line KL in [the supposedly resulting] triangle BDLKGB would be longer than straight line GD. It is assumed [above] to be equal, so GD is seen with more visual rays than KL. Therefore, GD is seen more distinctly than KL because [identical] things seen under more [visual] angles [i.e., by more rays] appear more distinct.

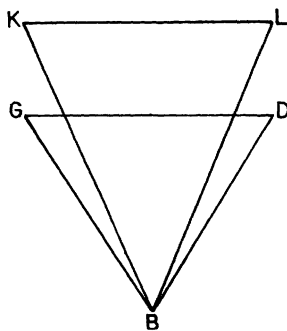


figure 2.1

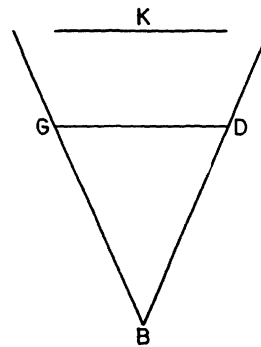


figure 2.2

[PROPOSITION 3] *Once a visible object reaches a certain distance [from the eye], it can no longer be seen.*

Let B [in figure 2.2] be the eye and GD a visible object. I say that GD can reach a point at which it will no longer be seen. For let it be placed at K in between the visual rays. Accordingly, none of the visual rays from B will reach K. But whatever the visual rays do not touch will not be seen. Hence, there is a certain distance at which any given object will no longer be seen.

Finding this explanation unacceptable, Ptolemy offers an alternative based on a dynamic conception of visual radiation: visual acuity is assumed to diminish with distance because of a continual weakening in the visual flux's intensity as it parts from its source at the vertex of the visual cone. Moreover, as will become clear in the following passage from Ptolemy's *Optics*, II, 20 (pp. 77-78), distance from the vertex of the visual cone is not the only factor affecting visual acuity; obliquity with respect to the visual axis also plays a role, visual acuity along the axis being optimal:

[2.4] And since [each] visual ray terminates at its own unique point, what is seen by the central ray—i.e., the one that lies upon the axis [of the visual cone]—should be seen more clearly than what is viewed to the sides [of the visual axis] by lateral rays. The reason is that those rays lie nearer to [the edge of the visual cone where there is an increasing] absence [of rays], whereas those rays that approach the [visual axis] lie farther from [such an area of] absence.⁴ The same holds for objects that lie toward the middle of spherical sections whose centerpoint is the apex of the visual cone, because the generating point of the sphere itself and powers that approach their generating sources are more effective. The farther such powers extend from their sources, then, the weaker they become—as, e.g., [the power of] projection [in relation to] the thrower, or of heat in relation to the heater, or of illumination in relation to the light-source. Therefore, since the visual ray within the cone has two primary referents, one being the centerpoint of the [ocular] sphere where the vertex of the visual cone lies and the other being the straight line that originates at this point and extends the whole length [of the cone] to form its axis, it necessarily follows that the visual perception of what lies far from the vertex of the cone is carried out by a more weakly-acting ray than the visual perception of something lying at a moderate distance. The same holds for objects that lie far from the visual axis in comparison to those that lie near it.⁵

Ptolemy thus treats the ray as a trajectory along which the visual flux hurtles like a projectile. But it is crucial to realize that this kinetic model is intended virtually, not really. In fact, as the following excerpt from *Optics*, II, 50 (pp. 91-92), makes abundantly clear, Ptolemy considers the ray itself as a virtual, not a real entity:

[2.5] It is not because it falls within the gap between visual rays that an extremely small visible object is not seen. On the contrary, it must be understood that, as far as visual sensation is concerned, the nature of visual radiation is perforce continuous rather than discrete. But if we set up mathematical demonstrations and treat the visual rays as if they are straight lines, [it follows that] large magnitudes lying the same distance away as small ones that are invisible because of their smallness will still be clearly seen. This

would not happen if the visual rays at that distance were diminished [in number] and spread out; instead, a large and a small object would look the same at that location. For if all the radiation that falls on their cross-sections along the entire base [of the visual cone] is composed of discrete rays, each of which apprehends a [single] point [on the given object], then, assuming that there is some real, spatial separation between those points, whatever lies far away ought not to be seen because the visual flux does not fall on those [interpunctal gaps]. Not even the points will be seen, because they have no size and do not subtend any angle. Hence, every such object [seen in this point-by-point way] will be invisible.

Taken as a discrete, mathematical line, the visual ray is obviously an analytic fiction for Ptolemy. The problem is how to justify the use of ray-analysis in accounting for intensity-variation; how, for instance, can we talk about “aggregation” and “concentration” of something that is continuous? Ptolemy’s response to this apparent contradiction is to treat radiation (whether visual or luminous) as virtually, not really discrete, with the flux or light being treated as particulate and the ray being reduced to a trajectory of sorts. As we shall later see, this kinetic model of radiation is particularly useful for the analysis of reflection and refraction.

2. *The Visual Perception of Physical Space*

Whether it is taken as a real entity or as a mere analytic fiction, the visual ray is expressly designed to account for the visual perception of physical space insofar as it mathematically defines the relationship between eye and external objects. What gives the ray such power of definition is its perfect rectilinearity, for which Hero gives a physical justification in *Catoptrics*, chapter 2 (p. 320):

[2.6] The arrows we see shot from bows may serve as an example. For, because of the impelling force the object in motion strives to move over the shortest possible distance, since it does not have the time for slower motion, that is, for motion over a longer trajectory. And so, because of its speed, the object tends to move over the shortest path.

Moreover, that the rays emitted by us travel at an immeasurable velocity is evidenced by the fact that, when we have closed our eyes and then reopened them to look at the heavens, it takes no perceptible time for [the visual rays] to reach the heavens, for just as our eyes are opened, we see the stars, even though the distance is, so to speak, infinite.

Hero’s justification for the rectilinearity of rays is thus contingent upon two interdependent suppositions. On the one hand, the radiation is so swift that it “does not have the time” to move more slowly. On the other hand, because it is pressed for time, it necessarily takes the shortest point-to-point path: i.e.,

a perfectly straight one. As we will see in the next section, Hero puts the least-path supposition to good use in accounting for the equal-angles law in reflection.

Perfect rectilinearity is not the only characteristic of visual rays. As Euclid points out in the fifth and sixth definitions of the *Optics* (p. 2), they also possess directional privilege. Hence, "objects seen with higher rays appear higher, whereas those seen with lower rays appear lower; and objects seen with right-hand rays appear to the right, whereas those seen with left-hand rays appear to the left." It is this directional privilege that enables the visual rays to map physical space within the plane defined by the visual cone's base. Propositions 10 and 12 of Euclid's *Optics* (pp. 16-18 and 20) illustrate the application of this principle:

[2.7] [PROPOSITION 10] *The more distant parts of planes lying below the eye appear higher.*

Let the eye at A [in figure 2.3] be placed higher than BED, and let visual rays AB, AE, AD, and AG make contact with it, AB forming the cathetus [i.e., being perpendicular] with the underlying plane. I say that GD appears higher than DE but that DE appears higher than BE.

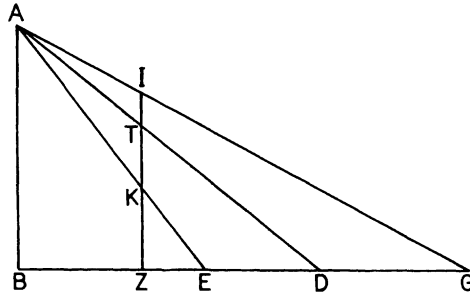


figure 2.3

Let point Z be taken on BE, and let perpendicular ZI be dropped. Because the visual rays reach ZI before ZG, let AG strike ZI at point I, AD at point T, and AE at point K. Thus, point I is higher than T, whereas T is higher than K. But G lies on the same ray as I, T on the same ray as D, and K on the same ray as E. But DG is seen according to AG and AD, whereas DE is seen according to AD and AE. Consequently, DG appears higher than DE. So, too, DE will appear higher than BE, for objects seen according to higher rays will appear higher.

[PROPOSITION 12] *Among magnitudes extending outward [beyond the eyes], those that lie to the right seem to verge toward the left, whereas those lying to the left seem to verge toward the right.*

Let AB and GD [in figure 2.4] be two magnitudes that are viewed, and

let E be the eye, from which rays ET, EK, EA, EZ, EI, and EG emanate. I say that EZ, EI, and EG seem to verge [successively] toward the left, while ET, EK, and EA seem to verge [successively] toward the right.

For, since EZ lies further to the right than EI, as does EI than EG, it follows that [the object seen by] EG appears to the left of [the object seen by] EI, as does [the object seen by] EI [in relation to the object seen] by EZ. By the same token, [the objects seen by] rays EK, EA, and ET appear to verge toward the right.⁶

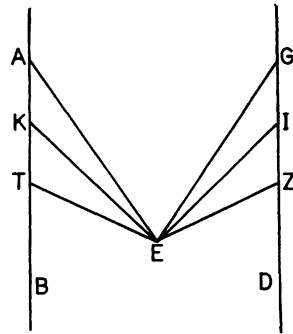


figure 2.4

Ptolemy, too, endows the rays with directional privilege in the plane of the visual field (i.e., up-down and right-left along the horizontal), but, unlike Euclid, he extends this directional privilege to the plane of the visual axis, thus allowing the visual faculty to judge distance along the vertical as well as relative disposition along the horizontal. Accordingly, as he summarizes it in *Optics*, II, 26 (pp. 81-82):

[2.8] The visual faculty also discerns the place of bodies and apprehends it by reference to the location of its own source-points [i.e., the vertices of the visual cones] . . . as well as by the arrangements of the visual ray falling from the eye upon those bodies. That is, longitudinal distance [is determined] by how far the rays extend outward from the vertex of the cone, whereas breadth and height [are determined] by the symmetrical displacement of the rays away from the visual axis. That is how differences in location are determined, for whatever is seen with a longer ray appears farther away, as long as the increase in [the ray's] length is sensible.

As far as Euclid is concerned, the cardinal determinant of spatial perception is the visual angle, which provides the basis for judging size. As he phrases it in the fourth definition of the *Optics* (p. 2), "objects viewed under a larger angle appear larger, whereas those viewed under a smaller angle appear smaller, and those viewed under equal angles appear the same size." Typical examples of this principle's application can be found in the first part of proposition 6 and in proposition 56 of Euclid's *Optics* (pp. 8-10 and 116):

[2.9] [PROPOSITION 6] *The intervals between parallels appear unequal when seen from an [increasing] distance.*

Let AB and GD [in figure 2.5] be two parallel magnitudes, and let E be the eye. I say that AB and GD do not appear [to maintain] the same distance apart, and that a nearer interval between them will always appear larger than a farther one.

Let visual rays EB , EZ , ET , EK , EI , and ED be dropped [to the respective parallels], and let BD , ZI , and TK be joined. Thus, since angle $BED >$ angle ZEI , BD appears longer than ZI . Moreover, since angle $ZEI >$ angle

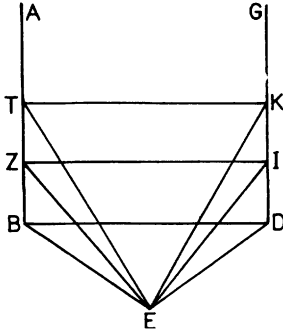


figure 2.5

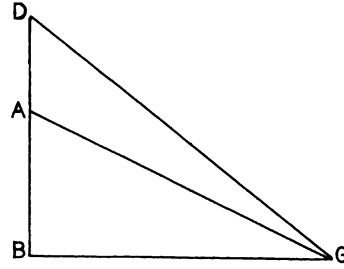


figure 2.6

TEK , ZI appears longer than TK . Therefore, BD appears longer than ZI , and ZI appears longer than TK . Consequently, parallels will not seem to lie the same distance apart [throughout] but will seem unequally separated.⁷

[PROPOSITION 56] *Magnitudes appear to get nearer as they get larger.*

Let AB [in figure 2.6] be the magnitude that will be seen, and G the eye, from which visual rays GA and GB emanate. Let BA enlarge to become BD , and let visual ray GD be dropped [to it]. Accordingly, angle $BGD >$ angle BGA , so that BD appears larger than BA . But identical objects that seem larger [than they are] appear to grow, and because objects closer to the eye appear larger, magnitudes will appear to the eye to get nearer as they enlarge.

Euclid is aware, however, that the visual angle alone is inadequate as a criterion for judging size. Distance is a critical factor as well; yet, as he demonstrates in proposition 8 of the *Optics* (pp. 14-16), the relationship between apparent size and distance is not a matter of simple proportionality (i.e., a given object seen at twice the distance will not appear half its original size):

[2.10] [PROPOSITION 8] *Equal and parallel magnitudes that lie at unequal distances from the eye do not appear proportional [in size] to the distances.*

Let AB and GD [in figure 2.7] be two [equal and parallel] magnitudes at unequal distances from the eye at E . I say that it is not the case that [the apparent size of] GD [in comparison to that of] AB is according to [the ratio of] BE to ED . For let the two rays AE and EG fall [to the respective magnitudes], and from centerpoint E at distance EZ let arc IZT be described. Therefore, since triangle $EZG >$ sector EZI , whereas triangle EZD

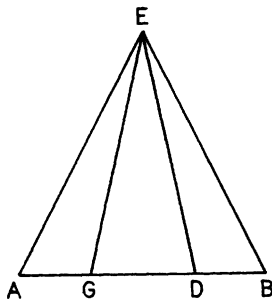


figure 2.8

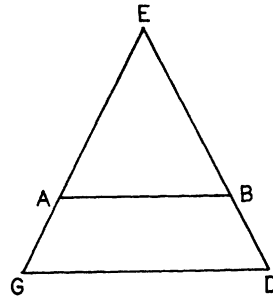


figure 2.9

And if there is a difference in [either] of the two remaining features alone, then the object will never appear larger, no matter whether it faces us more directly or whether it is closer. Thus, it will appear either smaller or equal, and in each case the apparent size will depend on relative differences in actual size.

[EXAMPLE II.1] For instance, if two magnitudes, **AB** and **GD** [in figure 2.9], have the same orientation and subtend the same angle at **E**, then, since **AB** does not lie the same distance as **GD** [from point **E**] but is closer to it, **AB** will never appear larger than **GD**, as seems appropriate from its proximity [to **E**]. Instead, it will either appear smaller (which happens when the distance of one from the other is perceptible), or it will appear equal (which happens when the difference in [relative] distance is imperceptible).

[EXAMPLE II.2] Likewise, if there are two magnitudes, such as **AB** and **GD** [in figure 2.10], that subtend the same angle at **E** and lie the same distance from it, and if their orientation is different, [so that] one of them, **AB**, faces **E** directly and the other obliquely, then **AB** will never appear larger than **GD** on account of its facing orientation. Instead, **AB** will either appear smaller than **GD** (which happens when the difference in orientation between the two magnitudes is perceptible), or it will appear equal to **GD** when the difference in orientation is imperceptible.

It seems, moreover, that size-comparison among these objects ordinarily springs from judgment rather than from the actual nature of the orientation or distance.

For, if a given sensible impression of the angles is aroused, if certain objects appear oblique or remote, and if one judges any of them to be smaller—even though it may be the only one—still, while one object may not be visually

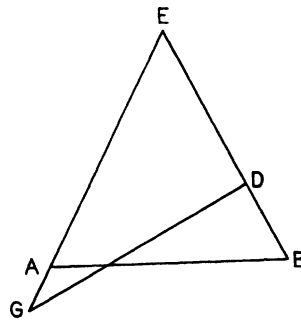


figure 2.10

sensed as smaller than another, but as equal [to it], we judge one of them to be larger. If the difference in obliquity or distance is imperceptible, then the two objects appear equal, but if that difference is perceptible, then they do not.

[EXAMPLE II.3] For example, in the [earlier] figure [i.e., 2.9], where the orientation was the same, if we draw the small angle formed by lines HZE and ETK [in figure 2.11], then magnitude GD will always appear larger than ZT, because it is farther away, and the angle it subtends is larger. But HK will never appear larger than AB, since the judgment based on the angle is not outweighed by a judgment based on distance alone.⁹ However, HK will appear smaller than AB if the distances and the angles differ perceptibly. But when they differ imperceptibly, the magnitudes will appear equal, just as in the first case.

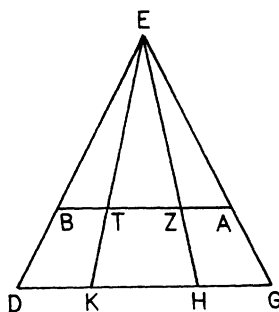


figure 2.11

[EXAMPLE II.4] On the other hand, in the [previous] figure [i.e., 2.10], where the distance was the same but the orientation was different, let us construct the small angle formed by lines KTE and EZH [in figure 2.12]. Accordingly, magnitude GD will always appear larger than ZK, for the size of [its subtended] angle and the [amount of its] obliquity conspire to make it appear larger. Yet HT will never appear larger than AB, because the judgment based on the angle is not outweighed by the judgment based on orientation alone.¹⁰ HT will, however, appear smaller [than AB] if the obliquity and angles differ perceptibly, whereas, if they differ imperceptibly, it will appear equal to it.

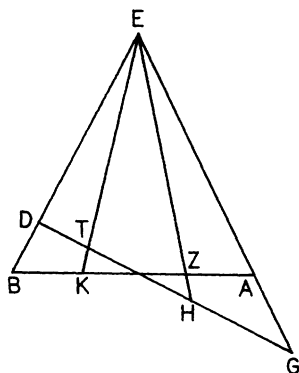


figure 2.12

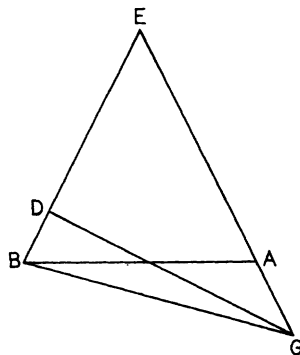


figure 2.13

[EXAMPLE II.5] If we leave both magnitudes oriented as they are in the preceding figure and connect line GB [in figure 2.13], then GB will always

appear larger than **AB**, since it is farther away and more oblique, and since there is no difference in the angles. Moreover, **GB** will appear larger than **GD** when the judgment based on distance outweighs that based on obliquity. But it will appear smaller than **GD** when the judgment based on distance is outweighed by that based on obliquity. Finally, they will appear equal when the judgments based on both variables are equal to the sense.

[EXAMPLE II.6] If, however, both magnitudes differ in all three previously cited respects, as happens when we construct the figure [2.14] with both magnitudes similarly oriented and draw line **GLMB**, then **GB** will always appear larger than **ZT**, because all three variables (i.e., size of angle, amount of obliquity, and distance) conspire to make it appear larger. It also appears larger than **HK**, because **GB** has two of the variables (i.e., size of angle and obliquity) that make an object [look] larger. On the other hand, **LM** always [appears] smaller than **GD**, since two variables make it do so—i.e., smallness of angle and proximity [to **E**]. The only variable belonging to **GD** to make it appear smaller is its facing orientation.

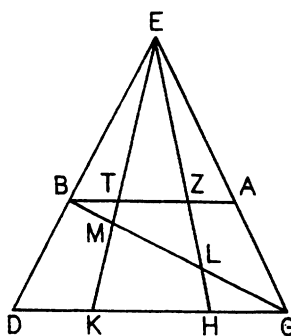


figure 2.14

Nevertheless, **LM** does not appear smaller than **AB** in all cases, for **AB** has nothing but the size of [its] angle to make it appear larger, whereas **LM** has greater distance and obliquity in its favor. Still, **LM** appears larger when the difference in these two features together outweighs the difference in angles. If the difference in these two features [together] is outweighed [by the difference in angles], though, **LM** will appear smaller than **AB**, and if the differences balance out, the magnitudes will appear equal.

That, then, is why we should not consider as adequate [the explanation of] those who have claimed that [in size-judgment] only distances should be taken into account along with angles and their resulting impression, whereas, they have claimed, any difference in orientation should be nullified in cases where the distances are equal. For, although there may be no difference in distance, a difference in orientation can still frequently occur. From this comes a perception of size different from the one that is due to the angle, as long, again, as we suppose this difference to be perceptible. Generally, therefore, since it seems that the size of visible objects is perceived by means of the size of the visual angles, that [angular] variable should be more fundamental and apposite [than the others in size-judgment].

Suffice it to say, because the interplay of these factors is complex and variable, it is not clear under precisely what circumstances one outweighs the other. Nonetheless, the very attempt to bring all three factors into account reveals

the sophistication of Ptolemy's approach to size-perception in comparison to that of Euclid.

So far we have accounted for the perception of location and size—two of the five “common sensibles” listed by Ptolemy at the beginning of the second book of the *Optics* (see passage 1.28 above). Of the remaining three (shape, motion, and rest), shape is discerned by the grasp that the impinging visual flux has of the object's boundaries. In the frontal plane of the visual field, shape is apprehended according to the cross-section of the visible object. Thus, for instance, a circular object viewed frontally will be adjudged circular because all of its diameters appear equal. Euclid makes this point in the first part of proposition 34 of his *Optics* (pp. 60-62):

[2.12] [PROPOSITION 34] *If a straight line is erected from the center of a circle at right angles to the plane of the circle, and if the eye is placed on [the endpoint of] that line, then all the diameters drawn in the plane of the circle will appear equal.*

Let point A [in figure 2.15] be the center of the circle, from that point let there be erected a perpendicular to the plane of the circle, and let the eye lie on that perpendicular at B. I say the diameters will appear equal. Let DG and EZ be two diameters, and connect BG, BE, BD, and BZ. Therefore, since [radius] ZA = [radius] AG, since AB is common, and since the angles [at A] are right, it follows that base BZ = base BG, and the angles formed at the [respective endpoints of those] bases are equal. Thus, the angle [ZBA] formed by ZB and BA = the angle [GBA] formed by AB and BG. Likewise, angle EBA = angle ABD. Accordingly, the angle formed by GB and BD = the angle formed by EB and BZ. But things seen under the same angle appear equal. Therefore, GD [appears] equal to EZ.

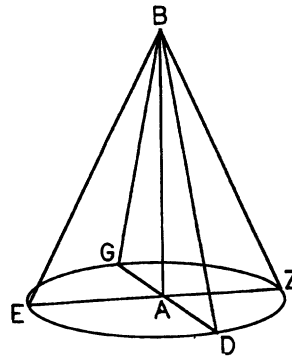


figure 2.15

Euclid then goes on to show in succeeding propositions that a circle viewed aslant does not appear circular but, rather, oblong. In all cases, however, the analysis is essentially two-dimensional: that is, it assumes that shape-discrimination is made solely on the basis of cross-sectional comparisons.

Ptolemy adds nothing of moment to the Euclidean account of shape-discrimination in the frontal plane (i.e., in two dimensions), but he does address the issue of shape-perception in three dimensions. As we have already seen in passage 1.36 above, color-contrasts provide one clue for such perception. Thus, according to the example Ptolemy gives there, an object appears convex if the

central portion of it is more brightly colored than its edges. As far as the mathematical analysis of shape-perception is concerned, though, Ptolemy offers the following account in *Optics*, II, 64-67 (pp. 98-99):

[2.13] The visual faculty perceives shapes by means of the shapes of the bases upon which the visual rays fall. And it perceives the outlines of the shapes by distinguishing the lines encompassing those bases as straight or circular, these being the two principles that differentiate shapes. But it perceives [the shape of] the entire defining surface [of the object]—in terms of whether it is plane or spherical—by means of the surface of the [visual cone's] whole base Now a surface is perceived [as] plane when straight lines fit everywhere throughout the entire base upon which the visual rays fall, whereas a surface is perceived [as] spherical[ly curved] when circular arcs fit everywhere on all portions of the base.

A surface or line appears concave to us when the ratio between the visual ray orthogonal to it and the oblique rays is greater than the ratio between the corresponding [perpendicular and oblique] rays falling on a plane surface or on [its constituent] straight lines. Such [surfaces or lines] appear convex to us when the [former] ratio is smaller [than the latter one]. These two shapes [i.e., concave and convex] are alike, though, [insofar as] they are proportional to one another and belong to the category of things, such as ascent and descent, that are deemed "reciprocal." Each of them is also congruent with the other, the convex with the concave and the concave with the convex, for what encompasses the outside is concave, whereas what is encompassed inside is convex. So too, objects are apprehended [as] concave by means of the surfaces of convex bases [defined by impinging visual rays], whereas objects are apprehended [as convex] by means of the surfaces of concave bases, just as such objects are perceived by touch, convex ones being apprehended through the concavity of the encircling hand, and concave ones being apprehended through the convexity of the encircled hand.

The perception of convexity and concavity is thus a function ultimately of distance-perception. In both cases, the shape of the surface is adjudged according to the comparative lengths of the central and outlying rays: when the central rays are proportionately shorter, the object is taken to be convex; when they are proportionately longer, it is taken to be concave. By the same token, an object is perceived as slanted when the rays touching one edge are sensed to be proportionately shorter than those touching the opposite edge.

As to motion and rest, finally, the correct perception of both is ultimately contingent on the viewer's innate sense of his own motion or rest. The centerpoint of the eye thus provides an absolute reference-point for judging motion or rest, but the Governing Faculty is the final arbiter, as the following passage from *Optics*, II, 76 (p. 103) makes clear:

[2.14] The particular way in which locomotion is perceived, however, re-

quires a special discussion if we are to explain how it happens and make sense of the sort of change it represents. Visual perception of the phenomena associated with locomotion depends primarily on the visual faculty itself [rather than on the object undergoing the change]. A particular object appears stationary either when the vertex of the visual cone undergoes no perceptible motion over a given continuous span of time, and the endpoint of a given ray apprehends sensibly one and the same spot on the object, or when it does undergo some motion, but it appears to the viewer that the displacement of the viewpoint and the displacement of the rays by which the object is visually grasped are equal during equal spans of time. The motion and rest of the visual flux, as well as their degree, is apprehended not by the visual faculty but by the sense of touch that extends to the Governing Faculty, in the same way that we do not discern the motion of our hands by sight when our eyes are closed but by means of a continuous [sense-link] that reaches to the Governing Faculty.¹¹ We do, however, see a given visible object move when the viewpoint [at the vertex] remains immobile over some minimal perceptible span of time while a noticeable change somehow occurs in the visual flux or when the displacement of the moving viewpoint is not equal to the displacement of the visible object vis-à-vis succeeding rays during equal spans of time.

Perception of motion and rest is therefore relative to the internal sense of self-motion or rest. If the viewpoint is stationary, for instance, and an object passes through the flux emitted from it, that object's motion is properly sensed. On the other hand, if the viewpoint is in motion without the viewer's sensing it (via the Governing Faculty), the visual flux emitted by him will share that motion, and, in passing over an external object, will cause the viewer to impute his own motion to that object (or he will deem it at rest if it keeps pace with the passage of the visual flux). Under normal circumstances, however, the viewer will sense his own motion and will accordingly discern the object's motion relative to it. Ptolemy illustrates these points in *Optics*, II, 78-81 (pp. 103-105):

[2.15] [EXAMPLE II.7] Take, first, the case of visible objects that move directly toward or away from the eye. Let **AB** [in figure 2.16] be a given visual ray, and let **A** be the vertex of the visual cone. Then, if the eye is stationary and does not move during some continuous span of time, and if the visual ray apprehends the visible object at point **B**, it will indeed appear stationary. If the object is apprehended elsewhere than at point **B**, it will of course appear farther away when it moves outward from point **B**, whereas it will appear nearer when it moves inward from point **B**.



figure 2.16

But if the eye moves directly toward or away from the object, e.g., from **A**

to **G** along line **AB**, and if the visible object **D** moves to **B**, then, if the place at which the visible object [**D**] is [finally] encountered is where point **B** previously lay, it will appear stationary, and ray **AB** will appear to have moved the same distance, **AG**, as the eye. If the visible object is encountered in some place other than **B**—if it is in fact at point **D**—then it will appear to move toward the eye at the same speed [as the eye is moving]. If, however, it is encountered at some point closer in than **D**, it will appear to move toward the eye but more quickly. And if it is encountered at some point between **B** and **D**, it will [appear to] move in the same direction but more slowly. Finally, if it is encountered anywhere beyond **B**, it will [appear to] move in the opposite direction.

[EXAMPLE II.8] Now let us take another example involving lateral motion. From point **A** [in figure 2.17], which is the vertex of the visual cone, let several rays, such as **AB**, **AG**, and **AD**, be drawn, and let them remain fixed, so they do not move. When the location of the visible object is the same—e.g., when it remains in line with **AG**—then it will appear stationary. But if its location is [subsequently] different—e.g., if it is [subsequently] in line with **AD** or **AB**—it will appear to move.

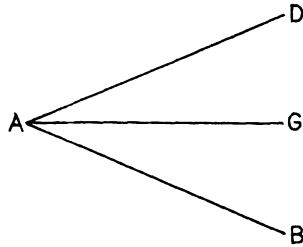


figure 2.17

On the other hand, if the visual cone pivots around its vertex **A**, and if we suppose that the distances between the given rays are equal throughout and that the motion occurs in the same [counterclockwise] direction, such that **B** [moves] to **G**, while **G** [moves] to **D**, then any [fixed] object that used to be seen along ray **AG** will be seen along ray **AB**, which is the next ray in line, and again that object will appear stationary, because the position of **AB** will be precisely the same as **AG**'s [used to be]. But if the object continues on to any of the rays other than this one, it will appear to move. For if it [continues] to be viewed along [moving] ray **AG**, its [new] position will be in line with **AD**, and it will appear to move in the same direction and at the same [angular] velocity as the eye. If it is seen by any of the rays ahead of **AG** [which will have moved to **AD**], it will appear to move in the same direction [as the eye] but faster. And if it is seen by any of the rays that lie between **AB** and **AG** [which will have moved to **AD** and **AG**, respectively] it will appear to move in the same direction [as the eye] but more slowly. Finally, if it is by any of the rays behind [**A**]**B** [which will have moved to **AG**] it will appear to move in the opposite direction.

The potential for misperception in such cases is extraordinarily high. Unless we sense our own motion, we are bound to misperceive either the direction or the speed of objects no matter whether they move in the frontal plane or the

plane of the axis. For that matter, we may even impute motion to the wrong object, as Ptolemy explains in *Optics*, II, 132 (p. 124):

[2.16] If we sail in a boat along the shore during twilight, or if we move in something other than a boat, and if we do not sense the motion of the thing carrying us, then we judge the trees and topographical feature of the shoreline to be moving. This illusion stems from the fact that, when the visual rays are displaced [laterally], we infer that the visible objects are moving because of the displacement of the visual ray. Although the visible objects are stationary, then, it is assumed that the apparent motion belongs to them.

Speed in relation to the distance and size of the moving object is also a factor in how we perceive motion. Certain threshold conditions obviously apply: a motion that is too fast or too slow under the given circumstances will not be properly perceived, as Ptolemy explains in *Optics*, II, 82 (pp. 105-106):

[2.17] Since a motion that occurs in the smallest perceptible span of time must be noticeable [in order to be perceived], we contend that if either of these [conditions] is not met, it is impossible for there to be noticeable motion. For, if the distance over which the motion occurs is moderate and the time within which it occurs is imperceptible (in fact, this often happens in circular motions, as is evident in the case of a potter's wheel or the wheels of horse-drawn chariots when they are violently turning), the motion is not apparent, because the [moving objects] return to their original positions in less than the minimum perceptible time-span while the visual flux remains fixed for what strikes the sense as a continuous time-span on the places that the visible object occupies—and that is what characterizes [the visual perception of] objects that do not move. On the other hand, if an unnoticeable motion occurs during a brief perceptible time-span (as frequently happens in the case of objects that move when there is a great distance between them and viewers—e.g., in the case of the bodies that revolve in the heavens or objects that [move] far off at sea), then [such objects] will not appear to move. And since every magnitude and every place is perceived by means of the distances that the visual flux senses, it is generally the case that, if the place to which a moving object makes its passage is imperceptible, neither the object nor its motion will be seen. And if this happens during a tiny perceptible [period of] time, we perceive neither when it occurs nor how long it took in terms of the [amount of] time we call “now.” Under these circumstances, then, nothing will appear [to happen] during a continuous span of time, but such objects will appear to move only if, by ignoring the sense-data, we plot their motion in another way on the basis of their original location.

As to relative distance, Ptolemy explains its effect in *Optics*, II, 100 (p. 111):

[2.18] Of objects that have the same speed, those that are closer appear to

move faster. Among those objects whose cross-sections are equal, the ones that are nearer to the viewpoint subtend a greater visual angle. Those things, moreover, that mark out greater arcs in equal times appear to travel more swiftly.

Finally, the case of relative size is addressed by Ptolemy in *Optics*, II, 101 (pp. 111-112):

[2.19] Even though they are smaller, [certain objects], such as small boats and small arrows, that are moving at the same rate as larger ones in fact appear to move more rapidly. Indeed, putting it another way, objects that traverse [incremental] spaces longer than themselves in several stages during equal times are adjudged to move faster than those that traverse [incremental] spaces equal to their own length, since a small object measures a given distance several times more than a large one. It therefore follows that, when objects of unequal size travel at the same speed over given distances, the smaller ones appear to move in more stages. Accordingly, then, since the object passes over the distance apprehended by the visual flux in equal times in more stages than [does the other object passing] over the same distance, it appears to move faster.

It is thus because of relative size rather than actual speed that small animals, such as mice, appear to dart, whereas larger (and often faster) animals, such as dogs, appear to do no more than lope.

3. *Binocular Vision*

The analysis to this point has dealt with monocular vision. Surprisingly enough, until Ptolemy's time, binocular vision seems to have been all but ignored, although it does figure obliquely in a series of six propositions (23-28) in Euclid's *Optics*. Here we will take propositions 23 and 26 (pp. 36-38 and 42-44) as examples:

[2.20] [PROPOSITION 23] *However it is viewed by a single eye, a sphere always appears smaller than a hemisphere, and the part of the sphere that is seen looks like a circle.*

Let there be a sphere with center **A** [in figure 2.18], and let **B** be the eye. Then let **AB** be connected, and let a plane be extended along **BA** so that it cuts a circle [through the sphere]. Let that circle be **GDTI**. Let circle **GBD** be constructed upon diameter **AB**, and let **GB**, **BD**, **AD**, and **AG** be drawn. Since **AGB** is a semicircle, then angle **AGB** is right, and so is **ADB**. Thus, straight lines **GB** and **BD** are tangent [to circle **GDTI**]. Let **GD** be connected, and let **IT** be drawn through point **A** parallel to straight line **GD**. The angles at **K** are therefore right. Now, if **AB** remains fixed [as an axis], and triangle **BGK** is rotated about right angle **K** until it returns to its original position, **BG** will touch the sphere at one particular point [throughout

The implication of this second theorem is clear: the larger image seen with both eyes is a composite of the two smaller images apprehended by each eye. But how are those two images fused into one coherent image? While Euclid has nothing to say on this subject, Ptolemy essays an explanation in *Optics*, II, 27-28 (pp. 82-83):

[2.21] Furthermore, every body appears at a single location to the [left or right] sides of anyone who looks with one eye alone, and it is seen at a single location by anyone looking with both eyes if he apprehends it with rays that are correspondingly arranged—that is, rays that have an identical and equal position within both visual cones in relation to their own axis. This happens when the axes of both cones converge at a given [spot on a] visible object, as is the case when we fix on visible objects with the normal glance that, by nature, lets us scrutinize them carefully.

It seems, moreover, that nature has doubled our eyes so that we may see more clearly and so that our vision may be regular and definite. We are naturally disposed to turn our raised eyes unconsciously in various directions with a remarkable and accurate motion, until both axes converge on the middle of a visible object, and both cones form a single base upon the visible object they touch; and [that base] is composed of all the correspondingly arranged rays [within the separate visual cones].

Under normal visual circumstances, then, the two eyes create a unified image by causing the two visual axes to focus on a single point on the given visible object, the result being that both visual cones share precisely the same base. However, as Ptolemy continues in *Optics*, II, 29-37 (pp. 83-86), anomalous visual effects can be created when the two axes are forced from their customary convergence—i.e., in the case of diplopia:

[2.22] But if we somehow force our sight from its accustomed focus and shift it to another object than the one we wanted to see, and if the [new] object toward which our sight is directed is somewhat narrower than the distance between our eyes, and if the visual rays [that] fall together from our eyes on that [new] object are not correspondingly arranged, then that same object will be seen at two places. But when we close or cover either of our eyes, then the image in one of the two locations will immediately disappear, while the other will persist, [and the image that persists is] sometimes the one directly in front of the covered eye and sometimes the one directly in front of the other eye. This point will be easily understood if we try to explain it in the following way:

Let a short ruler be set up, let two long, thin cylindrical pegs be stood vertically upon it, and let the distance between the two pegs themselves and between the pegs and the edges of the ruler be moderate. Let either edge of the ruler be placed between the eyes so that the pegs lie in a straight line at right angles to the line connecting the eyes.

Accordingly, if we focus our eyes on the nearer of the pegs, we will see it as one, whereas we will see the other, which is farther away, doubled. And if we close either eye, the peg that appears directly in front of that same eye and that forms one of the doubled images will disappear. On the other hand, if we focus our eyes on the farther of the pegs, we will see it as one and will see the nearer one doubled. And if we close either eye, the one of the doubled pegs that appears to be opposite to this same eye and that forms one of the two images will disappear.

That each of these phenomena occurs and is a logical consequence of what we have proposed will be seen from the following diagram:

[EXPERIMENT II.1] Let points **A** and **B** [in figure 2.20] be the vertices of the visual cones, and let **B** lie at the right eye and **A** at the left. Let two pegs, **G** and **D**, be erected vertically upon line **GD**, which is perpendicular to **AB**, and from each vertex of the two visual cones let rays **GA**, **GB**, **DA**, and **DB** be extended to the two pegs. Then let us first focus on **G**, which is nearer.

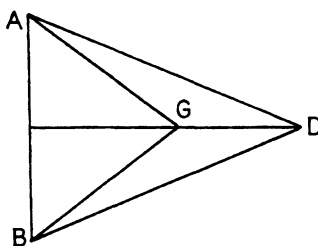


figure 2.20

AG and **BG** will therefore lie upon the axes themselves. Of the remaining two rays, however, **AD** will be one of the left-hand rays [in the cone whose vertex is at **A**], and it is obvious that **BD** is one of the right-hand rays [in the cone whose vertex is at **B**]. It necessarily follows, then, that **G** is seen at one location, insofar as each of the axes corresponds with the other. On the other hand, **D** must be seen at two locations, since **AD** is one of the left-hand rays of the left eye, while ray **BD** is one of the right-hand rays of the right eye. Thus, when we cover the left eye, the left-hand [member of the doubled image] will disappear, and when we cover the right eye, the right-hand [member] will disappear.

Now if we focus on **D**, the opposite will happen. This is demonstrated from the fact that **AD** and **BD** will lie on the [visual] axes. Thus, **D** will be seen as one, and **G** will be seen double, since **AG** happens to be one of the right-hand rays of the left eye, while **BG** is one of the left-hand rays of the right eye. So the opposite of what happened before will occur: that is, if we cover the left eye, the image seen on the right-hand side by ray **AG** will disappear, and if we cover the right eye, the image seen on the left-hand side by ray **BG** will disappear.

[EXPERIMENT II.2] However, if our focus is such that both axes do not converge on a visible object but are to all appearances parallel, like **AG** and **BD** [in figure 2.21], then each of the pegs will appear double according to the principles we have outlined.

For this to be manifest and demonstrable, the nearer peg, which lies at **L**,

should be white, while the farther peg, which lies at **M**, should be black. Points **L** and **M** will thus [each] be seen at two locations flanking the [actual] locations of **L** and **M**. Accordingly, if we cover the left eye, [both] pegs that [appear to] lie on the right-hand side will disappear, whereas, if we cover the right eye, [both] pegs that [appear to] lie on the left-hand side will disappear. The reason is that, when rays **AL**, **AM**, **BL**, and **BM** are extended, **AL** will certainly lie farther to the right [of axis **AG**] than **AM**, and **BL** will lie farther to the left [of axis **BD**] than **BM**. As a result, the right-hand [images of the] pegs will be seen by the left eye and the left-hand [images of the] cylinders by the right eye.

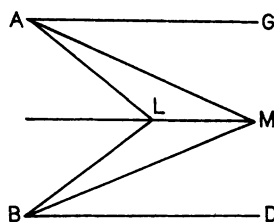


figure 2.21

So far, Ptolemy's analysis of diplopia has involved point-objects, represented by pegs. In the third book of the *Optics* he extends the analysis to lines, which are intended to represent any object with true spatial dimensions. The following example comes from *Optics*, III, 26-32 (pp. 141-143):

[2.23] [EXAMPLE III.1] Let us first discuss the case of convergence in which points **A** and **B** [in figure 2.22] represent the vertices of the two visual cones. Let line **AB** be joined, and let it be bisected at point **G**. Then, from that point let **GD** be drawn perpendicular to line **AB**. Let axes **AD** and **BD** converge at point **D**, and let the visible object be placed at **D**. Accordingly, this object will appear single and in its actual location.

So, too, if we draw line **EDZ** through point **D** perpendicular to **GD**, any object placed on this line will appear single and at its true location as long as it is aligned with point **D**.¹²

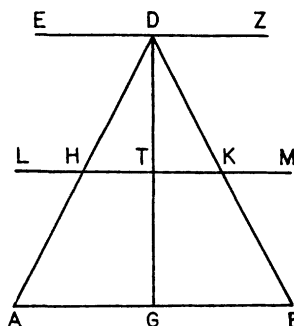


figure 2.22

On the other hand, if line **HTK** is drawn parallel to line **EDZ** while the two axes remain focused on **D**, an object at point **T** will appear at the two locations represented by **H** and **K**. Moreover, two magnitudes placed at those two locations [**H** and **K**] will appear at the three locations represented by points **T**, **L**, and **M**. At point **T**, of course, both will be seen together at once, as if they constituted a single object. They will be seen separately, however, insofar as **H** will appear at point **L** and **K** at point **M**, and each of the segments **LT** and **TM** will appear equal to **HK**.¹³

But if we focus both axes upon point **T**, we will see point **D** at points **E** and **Z**.¹⁴

That the number of these images proves to be as we have claimed can be

determined experimentally by anyone using the ruler with two pegs standing on it. Whoever wants to ascertain their true locations will do so by placing a finger on the visible object. For the finger actually touches the object when it appears in its true location, whereas, when it does not, the finger does not touch it but withdraws having encountered nothing there.

[THEOREM III.1] We have explained elsewhere the reason for the anomaly that we just discussed, and we have demonstrated it in a variety of ways when we dealt with objects that are seen at once with both eyes. Indeed, those objects that are seen with correspondingly arranged rays appear to lie in a single place even if there are actually two of them. However, if they are seen by means of non-correspondingly arranged rays, such objects, even if they are actually one object, appear to lie in two different locations.

To explain, let us join lines AE, AZ, ZB, EB, TA, TB, BH, and AK in the preceding diagram [as represented in figure 2.23]. Accordingly each of the spots at E, D, and Z will appear [respectively] in one location, since AD and BD lie on the visual axes themselves, and the rays that fall on E and Z are correspondingly arranged, since AE corresponds to BE, while AZ corresponds to BZ. On the other hand, since AH and BK lie on the visual axes, H and K will appear [combined] at single location T, but since BH and AK do not correspond, H and K will be seen at points L and M. Finally, because rays AT and BT do not correspond, point T will appear at the two locations H and K.¹⁵

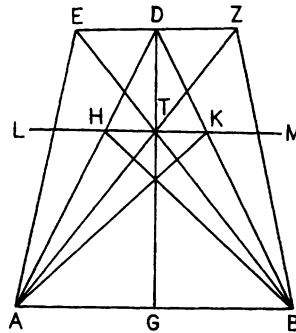


figure 2.23

To complete his analysis of binocular vision and diplopia, Ptolemy shows that what ultimately governs image-displacement in diplopia is the relative distance between the particular visual axes of both cones and what he calls the "common axis." His point will be clear from the following excerpt from *Optics*, III, 35-40 (pp. 144-146):

[2.24] Since all this is as we have claimed, it follows that nature equalizes the disjunction between the two visual axes and joins them according to the location of the visible object. Therefore, both of them fall upon the object from the Apex, which lies between them and is where the vertices of the visual cones ought to intersect.¹⁶ This Apex is equidistant from those axes, which have a common sensibility. And this is what keeps the vision tracking in a single direction midway between both sides, because it is impossible for

an object facing the visual axes to maintain the same orientation with respect to each of them. In fact, an object directly in front of both eyes is not orthogonal to both axes. How, then, can we determine spatial disposition if each of the axes is inclined to the other, unless, as we said, we do so on the basis of some intermediate reference-line whose distance [from the axes] is proportional [throughout]? Reason dictates that this reference-line be termed the "common axis."¹⁷

In the case of objects seen in direct vision, we need to provide a general determination of apparent location, and we should note that the displacement of such locations from the actual objects varies with the distance of the common axis from the proper visual axis. And this distance is [measured by] the perpendicular dropped from the visible objects to the common axis.

[EXAMPLE III.3] We must still consider whether the determining principle we just adduced actually agrees with observation. Accordingly, let us take the figure [for the first case in this series] as already drawn, so that each of its defining features maintains the same order [as in figure 2.24]. Let each of the lines AD and BD represent a visual axis. We will therefore find that whatever lies on line EDZ appears where it actually is, while whatever lies on line HTK will be encountered at places other than the one it [actually] occupies.

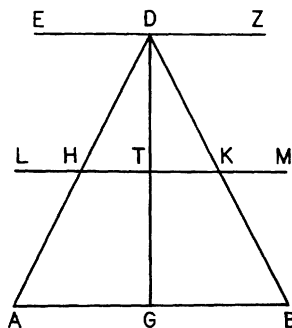


figure 2.24

Now it is clear that points E, D, and Z will appear in the places they actually occupy, because the perpendicular dropped from each of [the visual axes] to the common axis GD coincides with a single point. Since the displacement of the object's apparent location from its real location toward a given side will depend on the distance between the common axis and the proper axis, the visible objects themselves will be seen at their actual locations on account of the convergence of [all three] axes [at a single point]. It has therefore been demonstrated that each of those objects appears at its true location.

On the other hand, we have explained why H, T, and K do not appear in their true locations: the proper axis of the visual cone emanating from the eye at A is AH, and its displacement, measured by the perpendicular dropped to common axis GD, is HT, which is equal to line TK. Thus, the spot at which point T will be seen will be point K. That the same thing also happens if we look from the eye at B will be demonstrated in precisely the same way.

Likewise, if two objects are placed at H and K, and if LH and KM, respectively, are equal to HT and TK, then the object placed at H will appear to lie at the two points T and L. On the one hand, it will be seen at

point *T* from the eye at *A* along ray *AH*, because its perpendicular distance [from common axis *GD*] is *TH*; on the other hand, it will be seen at point *L* from the eye at *B*, because its distance is *TK*, the perpendicular linking the common axis and the proper axis *KB*, and *TK* is equal to *HL* and lies in a corresponding direction. This is the displacement of the apparent from the true location of *H* as seen from the eye at *B*. By the same token, the apparent location of the object placed at *K* will be at points *T* and *M*.

Perhaps the most glaring deficiency of Ptolemy's analysis of binocular vision has to do with his explanation of image-fusion. For if, as he supposes, this fusion is due to the two cones' sharing precisely the same base, the resulting composite and its two individual constituents must be identical among themselves. Contrary to the clear implications of the Euclidean theorem provided above, then, precisely the same image would be seen with two properly focused eyes as with one. As we shall see from the following excerpt, which is drawn from *On the Usefulness of the Parts* (X, 12, pp. 494-95), Galen recognized that the images would have to be disparate, at least as far as apparent location is concerned:

[2.25] Concerning [the fact] that an object is not seen in the same place by one eye as it is by the other, nor in the same place by both eyes together as it is by either one of them but is seen in one place by the right eye, in another by the left, and in still another by both eyes, this is what I must now say: Let the right pupil be at *A* [in figure 2.25], the left at *B*. Let the magnitude seen be *GD* and let visual rays from each pupil fall on *G* and on *D*, and when they have so fallen, let them be produced [to background object *TEIZ*]. By the right pupil the magnitude *GD* will be seen directly over against the magnitude *EZ*, by the left pupil it will be seen directly over against the magnitude *IT*, but by both directly over against *IE*. Hence neither pupil will see the object in the place where the other sees it, and both together will not see it where either sees it separately.

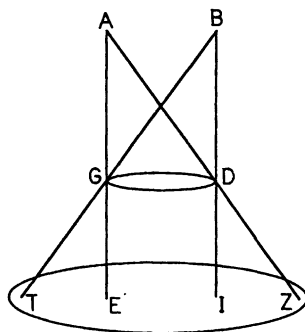


figure 2.25

Why did Ptolemy fail to recognize what Galen saw so clearly? Perhaps he simply failed to notice that the two eyes do not share precisely the same field of vision. But a likelier reason can be found in the model of image-fusion to which he was necessarily wedded. After all, to explain such fusion on strictly mathematical grounds, he had no choice but to account for it in terms of the

perfect symmetry of radial intersections at the visible surface. Otherwise, he would have been driven to conclude (as we do now) that image-fusion is, at bottom, psychologically rather than physically determined. Small wonder, therefore, that Ptolemy failed to understand the proper function of binocular vision in the perception of three-dimensional (rather than two-dimensional) space, a perception that allows us to perform such mundane, yet incredibly intricate tasks as catching a ball.

NOTES

¹In this case, Galen seems to be referring to the pupil itself, not the cornea.

²Unlike Ptolemy, Galen apparently locates the ocular source of visual radiation in the pupil rather than at the center of the eye itself.

³The Greek term for “visual ray” in this context is *opsis*. Elsewhere in the *Optics*, however, Euclid uses the term *aktis* to denote the same thing. Since the two are more or less interchangeable, I have made no effort to distinguish them in my translations.

⁴The rationale here is unclear. On the one hand, Ptolemy may simply mean that, since the visual radiation stops completely at the edge of the visual cone, rays that lie closer to that edge necessarily lie closer to the absence of radiation it demarcates. He could also be considering the radiation of visual flux from a central point in the eye to be ideally spherical in nature. Accordingly, the visual cone marks out a section in that sphere. Assuming that the radiation is uniform throughout the sphere, and assuming that the base of the visual cone (in external objects) is essentially planar, then the closer to the axis of the cone, the more concentrated the radiation, whereas the farther from that axis, the more dispersed (i.e., the closer to “absence”) it is.

⁵Overall, the point Ptolemy wants to make here is that objects are seen more clearly the closer they are (within certain limits) to the eye and the more closely the line of sight lies toward the axis of the visual cone. Note how vague Ptolemy’s formulation of the relationship between distance and intensity in light-radiation is; the implication, in fact, is that the relationship is a simple inverse one. It was only with Kepler that the proper inverse-square relationship was finally specified (for light rather than visual flux) in the early seventeenth century.

⁶The main point of this theorem, of course, is to explain the apparent convergence of parallel lines when they are viewed along a parallel line of sight within the same plane.

⁷Clearly, this theorem represents an alternative way of making the same point as proposition 12 excerpted above in passage 2.7.

⁸The point here, upon which Ptolemy elaborates later, is that the initial, primitive apprehension of size-differences depends upon the size of the subtended visual angles; distance and obliquity are ancillary variables that allow us to refine the data provided by those angles.

⁹According to Ptolemy, then, even if **HK** were longer than **AB**, it would never appear to be longer to the viewer at **E**, because the judgment based on visual angle takes precedence over that based on distance alone. What this means, of course, is that size-perception cannot be sufficiently explained on geometrical grounds.

¹⁰Here we are faced with basically the same inconsistency as adverted to above.

¹¹Note the obvious parallel between Ptolemy’s Governing Faculty and Galen’s “ruling principle” in passage 1.24 above.

¹²This is because the object on line **EDZ** forms a common base for both visual axes.

¹³The explanation for this set of claims is offered in the theorem that follows.

¹⁴Since the two axial rays **AT** and **BT** converge on it, point **T** will be seen at the center of the visual field. Point **D** will therefore appear to the left of that centerpoint when seen with ray **AD**, which is inclined to the left of axial ray **AT**, whereas it will appear to the right when seen with ray **BD**, which is inclined to the right of axial ray **BT**.

¹⁵In other words, as **H** shifts rightward to **T** for the eye at **B**, **T** shifts rightward to **K**, while **K** shifts rightward to **M**. Conversely, for the eye at **A**, as **K** shifts leftward to **T**, **T** shifts leftward to **H**, while **H** shifts leftward to **L**.

¹⁶The Apex (*principium*) seems to be the source from which the two visual axes are directed toward a common focus. That Ptolemy located it at the optic chiasma is certainly plausible but by no means unquestionable; see p. 37 above.

¹⁷This axis is therefore distinguished from what is termed the “proper axis” (i.e., the axis of the particular visual cone) in the next paragraph.

SECTION III

CATOPTRICS: ANALYSIS OF VISION BY REFLECTED RAYS

While Ptolemy's visual ray analysis of reflection yields many of the same conclusions as modern ray-analysis, there is one crucial difference. Modern ray-theory assumes that, because reflection is objectively determined, its attendant phenomena are more or less independent of subjective intervention—hence the division between “real” and “virtual” images. For visual ray theorists such an assumption is absurd, because it is the eye, not the object, that ultimately causes vision. *All* images generated by reflection are thus “virtual,” the product of subjective mediation rather than of objective, physical circumstances. Furthermore, since vision results from contact between the visual flux and visible objects, it follows that, whenever the flux reaches such objects after reflection, they must be seen. As we shall see, this imperative forces certain conclusions about image-formation that are both counterfactual and logically untenable by modern lights.

Our examination of reflection will start with a brief look at the equal-angles law and its theoretical and empirical justification. We will then turn to the problem of multiple reflections from concave mirrors, examining in some detail the conditions under which such reflections can occur and in what numbers. Next, we will address the basic principles determining image-location in plane, convex, and concave mirrors, after which we will deal with certain practical issues of image-formation and image-distortion in all three types of mirror. We will conclude with a cursory examination of visual effects produced by composite mirrors, our cardinal source in this case being Hero of Alexandria's *Catoptrics*.

1. *The Law of Equal Angles*

Since the goal of ancient mathematical optics was to explain the visual perception of physical space, the resulting theory of reflection was designed to account for image-formation, and deformation, in mirrors—the three basic mirrors chosen being plane, spherical convex, and spherical concave. The main principle governing image-location is of course the law of equal angles,

an early analysis of which is given in the passage below from proposition 1 of Euclid's *Catoptrics* (pp. 286-289):

[3.1] [PROPOSITION 1] *Visual rays are reflected at equal angles by plane, convex, and concave mirrors.*

Let B [in figure 3.1] be the eye and AC a plane mirror. Let visual ray BK emitted from the eye be reflected to D. I say that angle E = angle Z. Let perpendiculars BC and DA be dropped. Therefore, $BC : CK = DA : AK$, for this was postulated in the definitions.¹ Therefore, triangles BCK and DAK are similar. Accordingly, angle E = angle Z, because similar triangles are equiangular.

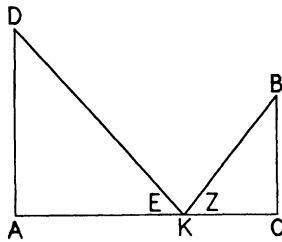


figure 3.1

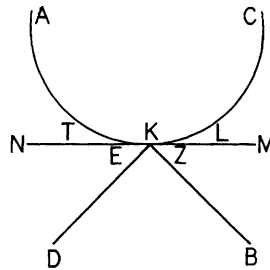


figure 3.2

Now let mirror AKC [in figure 3.2] be [spherical] convex, and let visual ray BK be reflected to D. I say that angle E + [horn] angle T = angle Z + [horn] angle L. I have applied plane mirror NM [to convex mirror AKC]. Accordingly, angle E = angle Z. But [horn] angle T = [horn] angle L, because MN is tangent. Thus, angle E + [horn] angle T = angle Z + [horn] angle L.

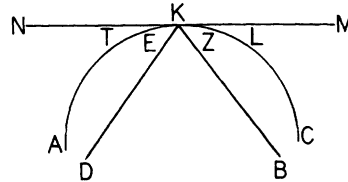


figure 3.3

Once again, let mirror AKC [in figure 3.3] be [spherical] concave, and let visual ray BK be reflected to D. I say that angle E = angle Z, for, with the plane mirror [NM] applied, [horn] angle T + angle E = [horn] angle L + angle Z. Therefore, the remaining angles E and Z are equal.

Unlike Euclid, Hero of Alexandria attempted a theoretical justification of the equal-angles law based on the dynamic model of radiation discussed earlier in passage 2.6 above. In reflection, the ray/projectile is blocked by an impermeable surface that forces it to rebound. Thus, as Hero puts it in the third chapter of the *Catoptrics* (pp. 322-24), "just as a rock hurled with great force against a compact body rebounds, . . . so also the rays sent forth by us at enormous velocity, as has been shown, are reflected by compact bodies when they strike them." On the basis of this dynamic model, Hero offers the fol-

lowing mathematical justification for the equal-angles law in chapters 4 and 5 of the *Catoptrics* (pp. 324-328):

[3.2] [CHAPTER 4] On the same grounds—that is, according to the speed of incidence and reflection—we will demonstrate that reflections occur at equal angles in plane and [spherical] convex mirrors. For again it is necessary to reason according to the actual minimum lines. I say, then, that of all incident and reflected rays reaching the same point from a plane or [spherical] convex mirror, those that form equal angles [with the reflecting surface] are the shortest possible. Moreover, if such is the case, they are reflected at equal angles according to an [underlying] rationale.

Accordingly, let AB [in figure 3.4] be a plane mirror, point G the eye, and D the visible [point-object]. Let GA be the incident ray, and let it be connected with [reflected ray] AD . And let the angle $EAG = \text{angle } BAD$. Then let a similar ray GB be the incident ray, and let it be connected with [reflected ray] BD . I say that $GA + AD < GB + BD$.

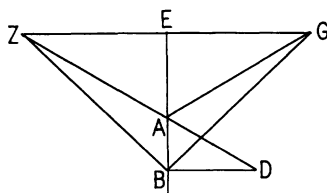


figure 3.4

For let GE be dropped from G perpendicular to AB , let GE and DA be extended to Z , and let ZB be joined. Angle $BAD = \text{angle } ZAE$, because they are vertical angles, and angle $BAD = \text{angle } EAG$ [which therefore = ZAE]. But the angles at E are right, so $ZA = AG$. Furthermore, $ZB = BG$, so that $ZD < ZB + BD$. But $ZA = AG$, and $ZB = BG$. Hence, $GA + AD < GB + BD$.

[CHAPTER 5] Let AB [in figure 3.5] be the surface of a [spherical] convex mirror, let G be the eye, and let D be the visible [object-point]. Let GA and AD form equal angles [with the reflecting surface], and let GB and BD form unequal angles. I say that $GA + AD < GB + BD$.

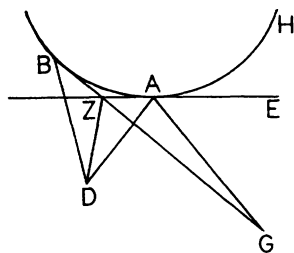


figure 3.5

For let tangent EAZ be drawn. Accordingly, [horn] angle $HAE = \text{[horn] angle } BAZ$, and the remaining [angle] $EAG = ZAD$. If, then, ZD is joined, it follows from the previous demonstration that $GA + AD < GZ + ZD$. But $GZ + ZD < GB + BD$, so $GA + AD < GB + BD$.

Taking an entirely different tack from Hero, Ptolemy attempts to establish the equal-angles law inductively. His expression of the law is somewhat different from Hero's in that he uses the normal to the point of reflection as the reference-line for the angles, so that "each of the ray's two branches joins at the

point of reflection and . . . both form equal angles with the normal dropped to that point" (*Optics*, III, 3, p. 131). Ptolemy's famous experimental confirmation of the equal-angles law is to be found in *Optics*, III, 8-11 (pp. 134-135):

[3.3] [EXPERIMENT III.1] Let a round, bronze plaque of moderate size, such as the one below [in figure 3.6], be set up, and let **A** be its center. Let both faces be planed down as carefully as possible, and let its edges be rounded and polished. Then let a small circle be inscribed at centerpoint **A** on either of its faces, and let it be **BGDE**. On this same face let two diameters, **BD** and **GE**, be inscribed to intersect at right angles; and let each quarter circle be divided into 90 degrees. Finally, let the two points **B** and **D** be taken as centerpoints, and, using **BA** and **DA** as radii, let the two arcs **ZAH** and **TAK** be inscribed.

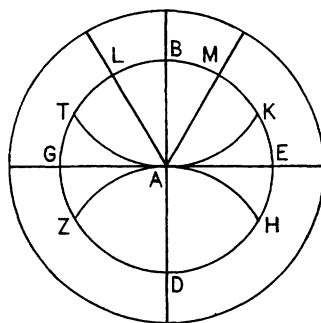


figure 3.6

Now let three thin, small, square, straight sheets of iron be formed. Let one of them remain straight, and let one of its sides be polished so that it appears as a clear mirror. Let the remaining two sheets be curved in such a way that the convex surface of the one and the concave surface of the other [taken together] form a circular section equal to circle **BGDE**, and let the two [respective convex and concave] surfaces of these sheets be polished so that they are made into two mirrors.

Let us cut arcs from each of the [above] two sheets, and let them be represented by **ZAH** and **TAK**. Let line **BA** be drawn in white and **AL** in some other color. Then let a small diopter be mounted upon **AL**, and let the aforementioned plaque be disposed in such a way that [the sight-line of] the diopter passes easily through point **L** and along line **AL**.² Now let the aforementioned plaque be placed with the side upon which the mirrors are [to be mounted] facing up. Of these mirrors, let the plane one lie on **GAE**, the convex one on **ZAH**, and the concave one on **TAK**. Finally, in the middle of the upper edge of each of the mirrors let a pin be attached axially to the mirror so as to keep it in place on point **A**.

Assume, then, that we view with either of our eyes through the diopter, which is placed at point **L** on **AL**, and that we direct our line of sight toward the axial pin of [each of] the mirrors. Accordingly, if we slide a small colored marker [along arc **TBK**] on the plaque's surface until it appears to us to lie on the same line of sight with **A**, then point **L**, point **A**, and the image that is seen in the three mirrors will appear [to lie] upon a single line [of sight].³ If, therefore, we mark the point at which the colored marker stands on the plaque's surface—i.e., the place from which the marker's image appears in those mirrors (e. g., the place represented by point **M**)—and if we draw out

straight line **AM**, we will find that arc **BM** is always equal to arc **BL**. Since that is the case, angle **LAB** will be equal to angle **MAB**, and line **BD** will be normal to all of those mirrors. Line **AL** [thus] delineates the [branch of the] visual ray incident to the mirror's surface from the eye, whereas line **AM** delineates the [branch of the] visual ray reflected from the mirror's surface to the visible object.

As we will see later, since Ptolemy bases his analysis of reflection from convex and concave mirrors on cylindrical sections, his conclusions about reflection from spherical mirrors are generalized from the simpler case of cylindrical mirrors.

2. Multiple Reflections and Multiple Images

Plane and Convex Mirrors: With the equal-angles law established, the next thing to determine is whether more than one ray-couple emanating from the same center of sight can reflect to the same object from a given center of sight. In other words, is it possible for the same object to project more than one image in a mirror? Excerpted below, proposition 4 of the Euclidean *Catoptrics* (pp. 292-294) addresses this issue for plane and convex mirrors:

[3.4] [PROPOSITION 4] *Visual rays that are reflected from plane and [spherical] convex mirrors will not intersect, nor will they be parallel.*

Let **AC** [in figure 3.7] be a plane mirror, **B** the eye, and **BCD** and **BAE** reflected ray-couples. I say that **CD** and **AE** are not parallel, nor do they intersect in the direction of **D** and **E**.

Now angle **Z** = angle **T**, and angle **K** = angle **M**. But angle **Z** > angle **K**, because it is an exterior angle of triangle **BAC**. Likewise, angle **T** > angle **M**. Therefore, **CD** is not parallel to straight line **AE**, nor does it intersect it in the direction of **E** and **D**.⁴

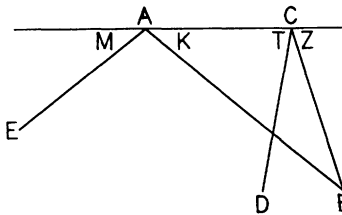


figure 3.7

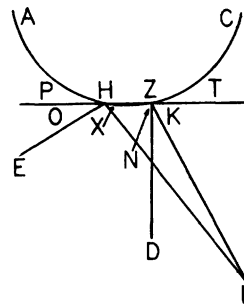


figure 3.8

This time let there be a convex mirror **AZC** [in figure 3.8], let **B** be the eye, and let **BZD** and **BHE** be reflecting ray-couples. I say that **ZD** and **EH**

are not parallel, nor do they intersect in the direction of E and D.

Let straight line HZ be drawn and extended out on both sides. Angle $K + \text{angle } T = \text{angle } N$ because the ray is reflected at equal angles. Thus, angle $N + \text{angle } AZH > \text{angle } K$. Moreover, angle $K > \text{angle } CHZ + \text{angle } X$, while angle $N + \text{angle } X > \text{angle } O + \text{angle } P$, for angle $X = \text{angle } O + \text{angle } P$. Therefore, angle $N + \text{angle } AZH > \text{angle } O + \text{angle } P$. *A fortiori*, then, angle $N + \text{angle } AZH > \text{angle } O$. Consequently, straight lines ZD and HE will neither meet nor be parallel.

Although the issue of multiple images is not explicitly addressed in this proposition, it nonetheless provides a clear focus. Since the separate reflected rays cannot meet, they will never reach the same object-point. Consequently, multiple images of that object-point cannot possibly be seen from the same viewpoint. Presumably, the additional restriction against parallelism is meant to cover cases of objects lying so far away from the mirror that, in reaching them, the reflected rays become virtually parallel.

Concave Mirrors: Having shown that multiple images are impossible in plane and convex mirrors, Euclid turns to the analysis of concave mirrors in proposition 6 of the *Catoptrics* (pp. 296-298):

[3.5] [PROPOSITION 6] *In the case of [spherical] concave mirrors, if you place the eye between the center [of curvature] and the surface, then the reflected visual rays will sometimes meet and sometimes not.*

Let AC [in figure 3.9] be a [spherical] concave mirror, D the centerpoint, and let the eye, B , be located between the centerpoint and the [reflecting] surface. Let visual rays BA and BC be reflected to I and Z [respectively], and let rays AT and CK be extended to the [continuation of the] mirror's surface. Accordingly, AT is either longer than, equal to, or shorter than CK .

Now if visual ray $AT = \text{visual ray } CK$, then arc $ACT = \text{arc } CTK$. Also, since angle $M = \text{angle } X$ —because angles of equal arcs are equal to each other—then angle $M + [\text{angle}] L = [\text{angle}] N + [\text{angle}] X$ according to [the equal-angles relationship in] reflection. Consequently, remaining angle $O = [\text{remaining angle}] P$, so that angle $R > \text{angle } O$. Now, since angle $R > \text{angle } P$, insofar as it is an exterior angle [of triangle BCR], and since $[\text{angle}] P = [\text{angle}] O$, then angle $R > [\text{angle}] O$. Let common angle ORZ be added [to angle O]. It follows, then, that CZ and AI will intersect in the direction of I and Z .⁵

The same will also hold if visual ray $AT > \text{visual ray } CK$. For angle $L + \text{angle } M > [\text{angle}] N + [\text{angle}] X$,⁶ $[\text{angle}] P > [\text{angle}] O$, and $[\text{angle}] R >$

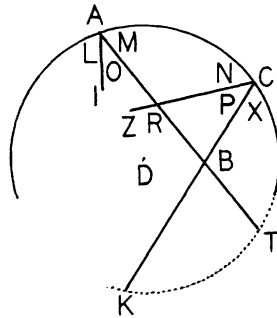


figure 3.9

[angle] O.

If $AT < CK$, then it follows that angle O > [angle] P. But [angle] R > [angle] P. Hence, nothing prevents [angle] R from being at least equal to [angle] O or less than it, in which case AI and CZ may not intersect.

As with proposition 4, so with this one, the issue of multiple reflections is not explicitly addressed but nonetheless provides the focus. That separate reflected rays can intersect at a given object-point therefore demonstrates that the same point can yield at least two images under certain conditions.

Repeated in both Hero's *Catoptrics* and Ptolemy's *Optics*, these theorems demonstrate that multiple images can be formed in spherical concave mirrors only. Suffice it to say, then, that the problem of image-formation and image-location for these particular mirrors is far more complex than it is for plane and spherical convex mirrors. In fact, as we shall now see, depending on where the object and eye are placed with respect to the reflecting surface, there is virtually no limit to the number of reflections that can occur in spherical concave mirrors.

Only three cases of multiple reflections are treated in the Euclidean *Catoptrics* and one in Hero's *Catoptrics*. The most general case, which is common to both Euclid and Hero, is covered in proposition 5 of the Euclidean *Catoptrics* (pp. 294-296):

[3.6] [PROPOSITION 5] *In concave mirrors, if you place the eye at the center [of curvature], on the [reflecting] surface, or beyond the surface (i.e., between the center [of curvature] and the surface), the reflected rays will intersect.*

Let ACD [in figure 3.10] be the concave mirror, and let B be the centerpoint of the [generating] sphere. Let the eye be placed at B, and let rays BA, BC, and BD strike the [reflecting] surface. Therefore, the angles formed at points A, D, and C are equal, for they lie on a semicircle. Thus, the reflected rays BA, BC, and BD will be diverted back upon themselves, for this point has been demonstrated [previously, in proposition 2]. They will therefore intersect at B.⁷

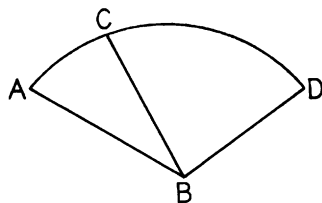


figure 3.10

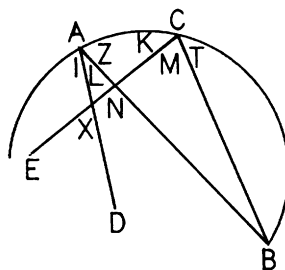


figure 3.11

Again, let ACB [in figure 3.11] be a concave mirror, and let the eye, B, be

placed on its surface. From **B** let reflected rays **BC** and **BA** strike at points **D** and **E**. Since arc **ACB** > arc **BC**, angle **Z** > angle **T**. Also, since [angle] **I** > angle **K**, [angle] **Z** + [angle] **I** > [angle] **T** + [angle] **K**. Since the remaining [angle] **L** < [angle] **M**, it is even smaller in relation to angle **N**. Thus, **CE** and **AD** will intersect at **X**. If the eye should fall beyond the surface, the same demonstration will hold, as in the following proposition.⁸

The remaining two cases are covered in proposition 6, which we have already examined in passage 3.5 above. According to the conditions set forth there, multiple images will invariably be seen when chord **AT**, containing incident ray **AB**, is either equal to or greater than chord **CK**, containing incident ray **BC**, because in both cases the reflected rays **AI** and **CZ** will intersect at a single object-point. If, however, **AT** < **CK**, and if the difference is great enough, the reflected rays will not intersect, in which case multiple images cannot be seen.

Whereas the Euclidean treatment of multiple images is sketchy at best, Ptolemy devotes considerable attention to the issue. As might be expected, his treatment is both more comprehensive and more systematic, involving an extensive case-by-case analysis of image-formation according to various dispositions of eye and object in relation to the reflecting surface. Accordingly, he asserts in *Optics*, IV, 2 (p. 175), "it is possible for one and the same visual ray to be reflected from a concave mirror [in several ways]: sometimes from all points on the mirror, sometimes from all points on one of its great circles, sometimes from three points, sometimes from two only, or sometimes from none." To illustrate Ptolemy's case-by-case approach to the problem, we will look at a selection of representative theorems. The first and simplest case is one with which we are already familiar: reflection to the centerpoint of the concave mirror. Ptolemy deals with this in *Optics*, IV, 3-5 (pp. 175-176):

[3.7] [THEOREM IV.1] Accordingly, we say that reflection from the entire surface of the mirror occurs as follows:

Let **ABG** [in figure 3.12] represent the arc of a circle lying on a concave mirror whose center is **D**. Let **BD** be drawn normal to the mirror. Let **D** be the center of sight and **E** the midpoint [where the axis of the visual cone intersects the surface] of the cornea. In this plane, on centerpoint **D**, let arc **ZEH** be drawn through the corneal surface. Then, let the two lines **AZD** and **GHD** be drawn. They will therefore be perpendicular to the mirror. Hence, all rays emanating from point **D** to arc **ABG** are reflected back on themselves to point **D**. Moreover, they are apprehended and sensed according to the spots at which rays like **DA**, **DB**, and **DG** strike [as they are] projected from the eye,

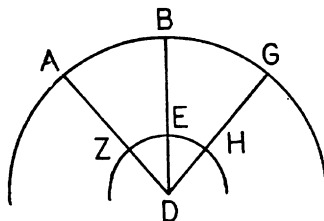


figure 3.12

and they are actually apprehended and sensed according to points E, Z, and H.

The same happens for all parts of the mirror that are sensed by the base of the visual cone. For if we assume that surface ADG forms a cone when it is rotated about axis BD, then what will appear in the whole surface of the mirror will be the image of the cornea, and it will appear continuous upon the entire surface. For the rays that are projected from point D to the mirror's surface and the normals that are reflected back upon them from the mirror meet at a point common to them all as well as to the reflecting surface, and this surface is the one that defines the locations of the image.

Although this may appear to be a mere rehearsal of the Euclidean demonstration in the first part of proposition 5 [passage 3.6] above, it is not. For one thing, Ptolemy extends the analysis beyond the great circle to the entire surface. For another, he is not content to represent the eye by a point but, instead, tailors his account to the fact that the entire ocular globe is involved. As a result, we can infer that he understood the eye to be a true sphere at whose precise centerpoint lies the source of visual flux. As far as visual sensation itself is concerned, finally, Ptolemy seems to have located it at the corneal surface, where the reflected rays "are actually apprehended and sensed according to points E, Z, and H" on the cornea. It is thus plausible to suppose that Ptolemy had *emphasis* in mind as the fundamental mechanism of visual sensation.

The second category of analysis for Ptolemy is defined by two conditions: the eye and the object must both lie upon a diameter of the sphere that defines the reflecting surface, and they must lie on either side of the center of curvature. Accordingly, our first example is from *Optics*, IV, 6-8 (pp. 176-177):

[3.8] [THEOREM IV.2] Reflection from [a single] circle on the circumference of such a mirror generally takes place according to the following example:

Let circle ABGD [in figure 3.13] be drawn with diameter BED. Let two points Z and H be marked upon BED on each side of centerpoint E, and let $EZ = EH$. Then, through point E, let diameter AEG pass orthogonally to diameter BED. Finally, let lines ZA and AH be drawn.

Thus, since $ZE = EH$, while line AE is common, and since the angles at E are right, angle ZAE = angle EAH. Hence, lines AH and AZ will be reflected at equal angles. Likewise, too, a reflection at equal

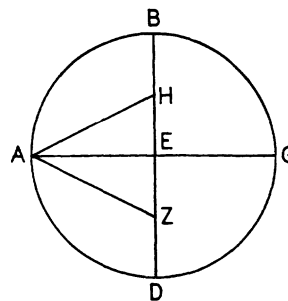


figure 3.13

angles occurs at point G. Now if we assume ABG to be the arc of a circle lying on a concave mirror, and if we suppose that surface AZH revolves about diameter BD , then points A and G certainly describe a circle on the mirror, and all the rays projected to that circle are reflected in the same way as they are along ZAH .⁹

Having thus analyzed the circumstances under which reflection takes place along a great circle of the sphere, Ptolemy turns to the analysis of reflection from a lesser circle parallel to that great circle. The following example is from *Optics*, IV, 11-13 (pp. 177-179):

[3.9] [THEOREM IV.5] We wish to demonstrate that on the arc between points A and B there is a point from which reflection takes place at equal angles. This demonstration is as follows:

Suppose [in figure 3.14] that the amount by which ZE exceeds EH (i.e., ZT) : $EH = ZH : KH$, KH being longer than HB . Let line KL be drawn tangent to circle ABG at point L , and let lines LZ and LH be connected. We therefore assert that HLZ reflects at equal angles.

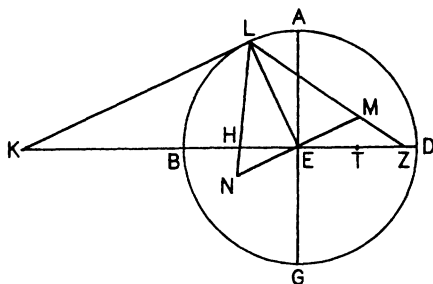


figure 3.14

Let line MEN pass through point E parallel to line KL , and let it meet the extension of line LH at N . Then let [normal] EL be joined, and cut off ET equal to EH . Therefore, since $ZH : KH = [ZT : EH =] ZT : ET$, then, by composition, KZ [i.e., $KH + ZH$] : $KH = ZE$ [i.e., $ZT + ET$] : ET , while $ET = EH$. By alternation, $KH : EH = KZ : ZE$. But $KH : EH = KL : EN$,¹⁰ and $KZ : EZ = KL : EM$. Thus, the ratio of lines EM and EN to one another is one of identity; so they are equal.¹¹ Moreover, line LE is common, and both angles of the two triangles $[ELN$ and $ELM]$ at E are right, because angle KLE is right. Hence, angles ELN and ELM will be equal.

Ptolemy then concludes his analysis of the second category (i.e., the case in which eye and object lie on the diameter) with the following proof from *Optics*, IV, 22-25 (pp. 181-182):

[3.10] [THEOREM IV.9] Accordingly, it will be shown as follows that, if reflection takes place at equal angles from some given point on a semicircle,

it is impossible for it to take place from any other point on that semicircle:

If it is possible, let the reflecting ray-couples be ZL , LH and ZS , SH [in figure 3.15], and let normals EL and ES be joined. Therefore, since both angles $[ZLH$ and $ZSH]$ are bisected, $ZL : LH = ZE : EH$ and $ZS : SH = ZE : EH$, so that $ZL : LH = ZS : SH$. By alternation, $ZS : ZL = SH : LH$, but line $ZS <$ line ZL ¹². Therefore, line $SH <$ line LH , which is false.

What we have said will also be demonstrated in like manner for any other semicircle, because, if the ray is reflected at equal angles from, say, point L , and if we assume that point B lies on a circular arc within the concave mirror and that surface HZL is revolved about diameter DB [as axis], then from every point on the circle formed by the circumscription of point L , the ray-couple connecting points Z and H will reflect at equal angles.

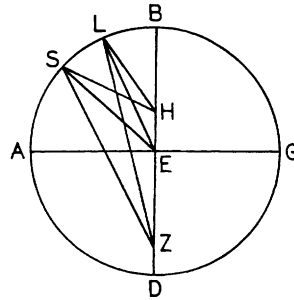


figure 3.15

Ptolemy's third category of analysis covers those cases in which the eye and object lie upon a chord located between the center of curvature and the surface from which the reflection takes place. In this case, eye and object are located on either side of the radius orthogonal to the chord. The following example is from *Optics*, IV, 26-30 and 34-35 (pp. 182-183 and 184-185):

[3.11] [THEOREM IV.10] After this, however, we ought to investigate the cases in which the appearances are created by three reflections from one arc while the eye lies at predetermined points.

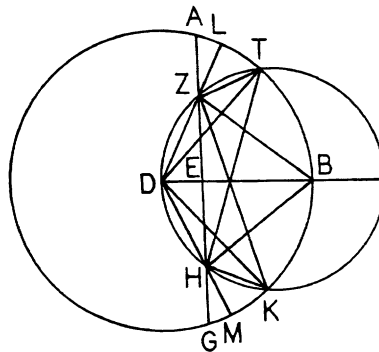


figure 3.16

Let ABG [in figure 3.16] be a circle with center D , and let AEG be drawn. Then let perpendicular DEB be drawn, and let arc ABG be marked off on a concave mirror. Let there be two points Z and H on line AEG such

[3.12] [THEOREM IV.15] In order not to prolong this discussion, we ought to stipulate that, if circle ZDH does not cut some arc other than BM , and if point Z passes through point A , as in the present figure [3.18], or beyond it [along line EA], then the reflections that occurred at equal angles from points B and K [in theorem IV.10, passage 3.11, above] will occur only from two places between B and K —as long as we exclude arc TL —because the whole of arc AB will lie within circle DZH .¹⁵ In the same way it will be demonstrated that none of the other rays reflects at equal angles.

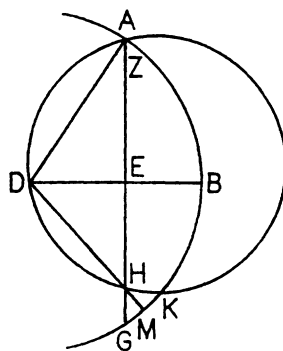


figure 3.18

[THEOREM IV.17] Nearly the same thing that we have shown also happens when circle ZHD does not cut circle ABG , and when point E lies between points Z and H . For in that case only one reflection takes place at equal angles.

Let circle DZH [in figure 3.19] not cut circle ABG , and, for a start, let $ZE = EH$.

It is obvious from what we have shown that there will be a reflection at equal angles from point B . Furthermore, it will be evident that there is no

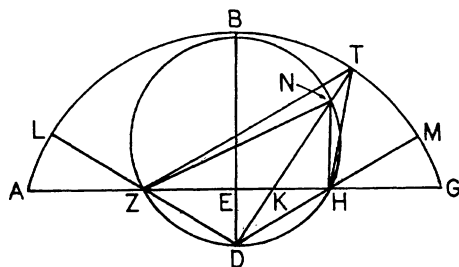


figure 3.19

reflection from arcs AL and GM . Now, it will be demonstrated as follows that it is impossible for a reflection [at equal angles] to occur from the arc that lies between L and B or the one that lies between B and M .

If such a reflection is possible, then let it take place according to ZT and TH . After [normal-]line $DKNT$ is drawn, lines ZN and NH will be joined. Since angle $ZND = \text{angle } DNH$, insofar as arc $DZ = \text{arc } DH$, it follows that angle $ZNT = \text{angle } HNT$. But angle ZTN was [supposed to be] equal to angle DTH . Therefore, triangles ZNT and THN are equiangular, and line TN is common. Consequently, $ZT = TH$, so it follows that $ZK = KH$. This entails a longer [line] being equal to a shorter [line], which is impossible.

The same thing will also be demonstrated if we assume the reflection to occur between points **L** and **B**.

[THEOREM IV.28] Now, let **ABG** [in figure 3.20] be a circle with centerpoint **D**, let line **AG** be drawn, and let **DEB** be dropped perpendicular to it. Let arc **ABG** be the segment of a circle inscribed in a concave mirror. Then, of the two points [eye and object] under discussion in the preceding figures, let one be placed at point **A** and the other at point **E**. We say, then, that there is no reflection at equal angles between them from arc **ABG**.

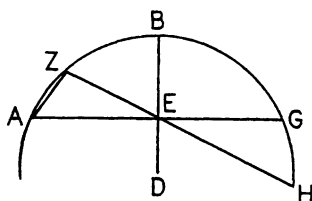


figure 3.20

It is clear that no such reflection occurs from arc **BG**, since it is impossible for the rays reflected between the aforesaid [two points] to encompass point **D**. From arc **AB**, as well, it is obvious that no reflection at equal angles can occur. For if such a reflection could occur, for instance, along **AZ** and **ZE**, then arc **HGZ** would be equal to arc **AZ**, which is false, because arc **HGZ** is much larger than arc **ZA**, given that arcs **AB** and **BG** are equal.

The final category of analysis deals with the case in which the eye and the object lie on a chord that is located beyond the center of curvature. Our example is taken from *Optics*, IV, 84-96 (pp. 199-201):

[3.13] [THEOREM IV.25] We must also demonstrate that, on the whole, when the center of the sphere lies between the mirror itself and the line joining the eye and the visible object, reflection at equal angles takes place from only one point.

[CASE 1]

Let **ABG** [figure 3.21] be the arc of a circle lying on a concave mirror with center **D**, and let normal **BDE** be drawn with line **HEZ** cutting it at right angles. Let one of the [end]points [of the reflecting ray-couple] be **Z** and the other **H**. First, let **EZ = EH**. We say, then, that between points **Z** and **H** there is no reflection at equal angles except from point **B**.

Let **BZ** and **BH** be joined. Accordingly, since lines **EB** and **EZ** are equal to lines **BE** and **EH**, respectively, and since the angles at **E** are right, the two triangles will be equiangular. Thus, angle **BZE** = angle **BHE** and angle **ZBE** = angle **HBE**.

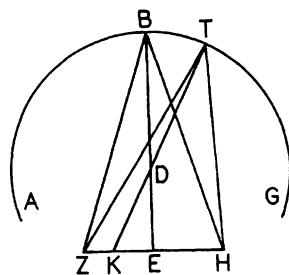


figure 3.21

If it is possible, let another ray-couple, such as **ZTH**, reflect at equal angles, and let normal **TDK** be drawn. Therefore, since line **KT** bisects angle **ZTH**, then $KH : KZ = TH : TZ$.¹⁶ But

$KH > KZ$, so $TH > TZ$. Thus, angle $TZH >$ angle THZ , and so angle BZH will be even greater than angle BHZ , which cannot possibly be the case, since we have just shown them to be equal. Consequently, TZ and TH do not reflect at equal angles.

What we have established will also be evident if we assume some other line.

[CASE 2]

Let a diagram [figure 3.22] be constructed similar to the previous one, but let $EH > EZ$. Then let lines ZDT and HDK be drawn, let line ZH be bisected at point L , and let angle ZDH be bisected by line DM . Let $EZ = EC$, and let lines LDN , MDS , and CDU be drawn. We say, then, that the reflection at equal angles of the ray-couple joining Z to H will occur between points N and S .

But from arc KA and arc GT no ray-couple will reflect at equal angles, for the ray-couple reflected from either of them does not encompass centerpoint D .¹⁷

So let us demonstrate first that it does not reflect from point S . If it could, let it reflect along ZSH . Therefore, since angle $ZDM =$ angle HDM , angle $SDZ =$ angle SDH . However, angle ZSD was [supposed] equal to angle DSH . Therefore, triangles ZDS and HDS will be equiangular, with line DS common. Accordingly, line $ZD =$ line DH , so angle $DZH =$ angle ZHD . But angle $DZH =$ angle DCZ , because sides CE and EZ are equal, as are the angles at E in triangles CED and EZD , and line ED is common. Consequently, in triangle CDH interior angle $ZHD =$ exterior angle DCZ , which is false.

If, moreover, we assume a reflection [at equal angles] from some point between S and T , the result becomes even more absurd. For in that case ZD ought to be longer than DH . Also, line MH , which is the longer, will be shorter than ZM , which is the shorter, and interior angle $DHZ >$ angle DZH , which is equal to exterior angle DCZ , all of which is impossible.

Again, assuming that it is possible, let the reflection occur along ZNH . Thus, since line $ZL =$ line LH , then $ZN = NH$, and angle $NZH =$ angle NHZ . It therefore follows that angle $NLZ =$ supplementary angle NLH , which is impossible.

This will be even more absurd, however, if we assume the reflection [at equal angles] to occur between K and N . Such an assumption leads to the conclusion that a line shorter than ZN is longer than a line longer than NH , while angle NLZ , which is more acute than angle NLH is larger than an angle that is more obtuse than angle NLH .¹⁸

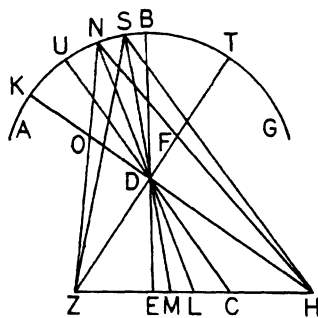


figure 3.22

[CASE 3]

But if there is a reflection at equal angles, it will certainly fall between points S and N, in which case nothing results that is contrary to what we have proposed. That only one ray-couple reflects at equal angles from this arc will be obvious, as we shall now explain. But in order not to confuse the figure, let us assume that the two points S and N lie on the arc from which we have claimed that it is possible for reflection [at equal angles] to occur. Moreover, assuming that it is possible, let the two ray-couples [ZSH and ZNH] be reflected at equal angles from those two points.

Accordingly, $ZS : SH = ZM : MH$, and $ZN : NH = ZL : LH$. But $ZL : LH > ZM : MH$, so $ZN : NH > ZS : SH$. By alternation, $ZN : ZS > NH : SH$, but $NH > SH$, because it is nearer to centerpoint D. Therefore, ZN will be much longer than ZS, which is impossible, since $ZS > ZN$, insofar as it is nearer the centerpoint.

But if we transpose N to point U and S to a position between points S and B, the arc from which the reflection occurs is defined by the line joining the centerpoint and the midpoint of the line joining the two points [Z and H] along with the line that bisects the angle subtended by this line [ZH]. And there is only one ray that reflects at equal angles from these points.

Although Ptolemy has been criticized by modern commentators for the cum-brousness of his case-by-case approach to reflection from concave mirrors, we should bear in mind that, compared to Euclid's extraordinarily spotty treatment, Ptolemy's is truly remarkable for its comprehensiveness as well as its structural integrity. Bear in mind, first, that Ptolemy's analysis proceeds in perfectly logical sequence according to the relative placement of object and eye with respect to the center of curvature. Specific cases of reflection depend upon whether: 1) eye and object coincide at the center of curvature itself; 2) eye and object lie on a diameter at various points with respect to the center of curvature; 3) eye and object lie on a chord between the center of curvature and the surface of reflection; 4) eye and object lie on a chord beyond the center of curvature. Furthermore, each of these cases is differentiated by subcase according to the number of reflections that can occur. Overall, then, Ptolemy's approach is as determinate as it is sensible.

3. *The Principles of Image-Location*

Although the equal-angles relationship between the angles of incidence and reflection is the most fundamental determinant of image-location, there are three other principles involved. Ptolemy sums them up in *Optics*, III, 3-5 (pp. 131-132):

[3.14] The first of these principles states that objects seen in mirrors appear

along the extension of the [incident] visual ray that reaches them through reflection, the resulting line-of-sight being determined by the placement of the pupil with respect to the mirror. The second principle asserts that particular spots [on a visible object] seen in mirrors appear on the perpendicular dropped from the visible object to the mirror's surface and passing through it [i.e., the cathetus of reflection] . . . A line normal to the surface of a sphere is commonly defined as one that forms right angles with all the tangents to the sphere at the common point on its surface [where the perpendicular is dropped]. Hence, it necessarily follows that all normals to the surface of spheres, when they are extended beyond, pass through the sphere's center.¹⁹

The implications of the principles just articulated will become evident according to observed facts, as we shall explain. For in all mirrors we find that, if we mark on the surface of any of them the spots where visible objects appear and then cover them, the image of the visible object will no longer appear. But afterward, if we uncover one spot after another and look toward the areas revealed, the designated spots will appear together with the image of the visible object along the line of sight projected from the vertex of the visual cone.²⁰ Also, if we stand long, straight objects at right angles to the surfaces of mirrors, and if the distance is moderate, their images will appear [to lie] perfectly in line with those objects [as they] are properly viewed outside the mirror.²¹

From both these instances it follows that a visible object must appear in the mirror at the point where the [extension of the incident] visual ray and the normal dropped from the visible object to the [surface of the] mirror intersect. [From this the third principle follows: namely, that] since they intersect, these two lines must lie in the same surface, and this same surface that contains them must be perpendicular to the mirror's surface, because one of the [contained] lines is perpendicular to the mirror's surface. Also, the [branch of the whole] visual ray that is reflected to the visible object will be in the surface just described. Finally, the normal dropped to the point of reflection on the mirror's surface must be the common section of all the various surfaces [of reflection] defined by the reflection of visual rays [at that particular point].

The four principles enunciated above can be easily understood by reference to figure 3.23: RI is the extension of incident ray ER along which the image is seen (principle 1); OI is the normal (i.e., cathetus of reflection) dropped from the object-point to the mirror's surface (principle 2); and EOI forms a single surface (the "surface of reflection") containing all the defining rays and normals (principle 3). Note, however, that in the case of convex or concave mirrors, the cathetus of reflection OI passes through the center of curvature and is perpendicular to the plane surface tangent to the surface of reflection at the point where it intersects that latter surface.

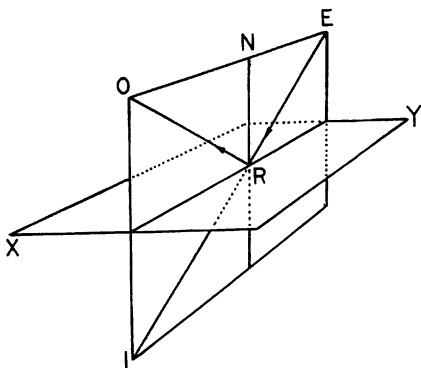


figure 3.23

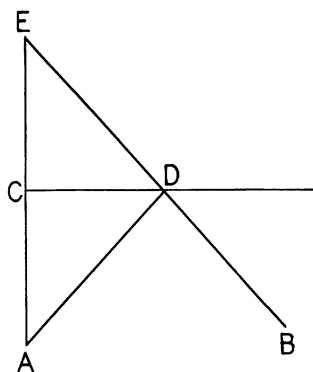


figure 3.24

For the Euclidean and Heronian versions of the first two principles—less elegantly articulated than Ptolemy's—we turn to proposition 16 of the Euclidean *Catoptrics* (p. 312) and chapter 6 of Hero's *Catoptrics* (p. 330) where the principles are laid down for plane mirrors:

[3.15] [PROPOSITION 16] *Everything that is seen in plane mirrors is seen along the straight perpendicular line dropped from the visible object.*

Let CD [in figure 3.24] be a plane mirror, B the eye, and A the visible object. From that object let perpendicular AC be dropped to the mirror. Therefore, since it has been adduced from experience that A is not seen to occupy point C,²² A will be seen along the extension of straight line AC. But it is also seen along the extension of visual ray BD, so it is seen at E. For we have supposed that a straight line is one whose intermediate points lie in the way of the endpoints,²³ so that AE and BE are rectilinear.

[CHAPTER 6] *In plane mirrors there is a spot where, when it is blocked, an image will no longer be seen.*

Let AG [in figure 3.25] be a plane mirror or an extension of its surface-plane. Let B be the eye and D the visible object, and let normals AD and BG be dropped to the mirror. Then let AG be cut at H such that $AD : BG = AH : HG$. I say, then, that when point H is blocked, D is no longer seen.

Let BH and HD be drawn. Hence, according to the proportionality [of sides given above], triangles [HAD and HGB] are similar. Angle E = angle Z, then, so D will appear through H. Hence, if that spot is blocked with wax or some other agent, D will no longer be seen. If, however, one removes the blockage

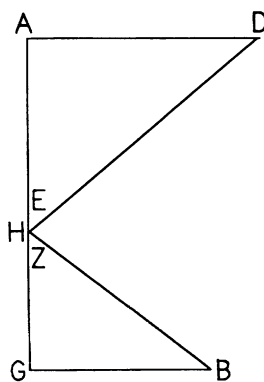


figure 3.25

at **H**, the image will appear in the mirror. For all rays incident upon a mirror will be reflected at equal angles.

Finally, in the Euclidean *Catoptrics*, the principles underlying image-location for spherical convex and spherical concave mirrors are provided in propositions 17 and 18 (pp. 312-14):

[3.16] [PROPOSITION 17] *Everything that is seen in [spherical] convex mirrors is seen along the straight line dropped from the visible object to the center of the sphere [i.e., the center of curvature].*

Let **CD** [in figure 3.26] be a [spherical] convex mirror, **B** the eye, and **BD** the visual ray reflected to the visible object **A**. Let **Z** be the center of the sphere, let [cathetus] **AZ** be connected, and let visual ray **BD** be extended to **E**. Therefore, since it has been adduced from experience that **A** is not seen to occupy point **C**,²⁴ **A** will be seen along the extension of straight line **AC**, where visual ray **BD** and [cathetus] **AZ** intersect: i.e., at point **E**, just as is the case in plane mirrors.

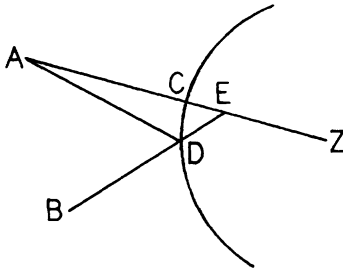


figure 3.26

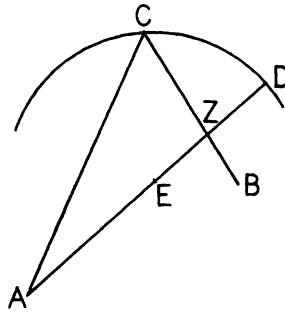


figure 3.27

[PROPOSITION 18] *Everything that is seen in [spherical] concave mirrors is seen along the straight line dropped from the visible object to the center of the sphere [i.e., the center of curvature].*

Let **CD** [in figure 3.27] be a [spherical] concave mirror and **BC** the visual ray reflected to the visible object **A**. Let **E** be the center of the sphere, and let a straight line be drawn from **A** to **E** and extended. Therefore, since it has been adduced from experience that **A** is not seen to occupy point **D** but must necessarily be seen along the extension of [cathetus] **AE**, it will be seen at the point where [cathetus] **AD** and visual ray **BC** intersect: i.e., at point **Z**.

In the case of plane and convex mirrors, all of the images generated are “virtual,” insofar as they appear behind the mirror, where the intersection of the cathetus of reflection and the extension of the incident ray must always lie. In the case of concave mirrors, however, the situation is considerably more com-

plex, for, depending on the relative position of the eye and object, the intersection that defines the image can: (1) be indeterminate (i.e., the incident ray and cathetus will never meet); (2) lie behind the mirror (creating a “virtual” image), (3) lie between the mirror and the center of sight (creating a “real” image), (4) lie at the center of sight itself, or (5) lie beyond the center of sight. Cases 1, 2, and 5 are discussed by Ptolemy in *Optics*, IV, 63-65 (pp. 192-193):

[3.17] [THEOREM IV.22] In the case where the line-segment reflected from the mirror to the visible object is equal to the line connecting the visible object and the center of the sphere, it necessarily follows that the [incident] ray-segment emanating from the eye to the mirror and the normal dropped from the visible object to the mirror are parallel [case 1 above]. But in the case where the two aforesaid lines are unequal, if the line connecting the center of the sphere and the visible object is longer than the reflected line-segment, then the intersection-point we have cited will lie behind the mirror [case 2 above]. If, however, the reflected line-segment is longer than the line joining the centerpoint [of the sphere] and the visible object, then the intersection-point will lie behind the eye [case 5 above].

Let ABG [in figure 3.28] represent the arc of a circle lying on a concave mirror with centerpoint D . With the eye placed at point E , let the reflection be assumed to occur at equal angles along EB and BZ . Then, from point D , let line DH be drawn equal to HB . Likewise, from point D , let line DT be drawn longer than TB , let line DZ be drawn shorter than ZB , and let H , Z , and T represent the visible objects. Finally, let each of the lines drawn from E to the visible objects fall, like EZ , between the center of the sphere and the mirror. We say, then, that DH is parallel to EB , that DT intersects it behind the mirror, and that DZ intersects it behind the viewer.

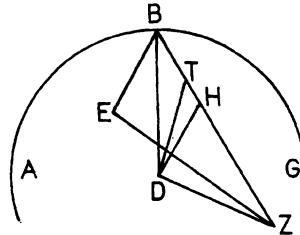


figure 3.28

Now, since $DH = BH$, angle DBH , which is equal to angle DBE , will be equal to angle BDH . Thus, angle $BDH = \text{angle } DBE$, and so line EB will be parallel to line DH . Again, since line $BT < DT$, angle $BDT < \text{angle } DBT$, which is equal to angle DBE . Thus, lines EB and DT intersect on the side of B and T , which lie beyond the mirror, since BZ is always reflected from the other side of line BD . By the same token, if angle $BDZ > \text{angle } DBZ$, which is equal to angle DBE , lines BE and DZ must meet on the side of points E and D , and it is evident that the intersection occurs behind the eye, because ZE lies between B and D .²⁵

Case 4, in which the image-point lies at the center of sight, occurs when line EZ coincides with cathetus DZ , which therefore intersects incident ray EB at

point E. Case 3, in which the image-point lies between the center of sight and the mirror, occurs when line EZ lies beyond cathetus DZ, which therefore intersects incident ray EB between E and B.

Precisely what is seen in the various cases described above? Ptolemy addresses this question in *Optics*, IV, 69-70 (pp. 194-195):

[3.18] But if . . . the image has no determinate location where it may be seen, then the sight fixes on a location common to both the image and the mirror. The same thing happens when the two straight lines just mentioned intersect in such a way that the image appears at the center of sight. But when the point of intersection lies behind the mirror, the image-locations preserve their proper disposition, for they appear behind the mirror. When, however, the distances [of such intersection-points] become inordinate, they are always [perceptually] shortened.

In addition, when the point of intersection lies behind the eye, the visible object is not seen behind the viewpoint, for such a perception is impossible. Rather, it appears in front of the mirror, even though the impression arising in the sense does not render the image in its proper place but transfers it to a location that is not its own, just as we have already pointed out. For the location, which is not properly assigned, renders the visible object on the very surface of the mirror. But when the location is properly determined, the image appears where it ought to according to the eye. And this is how images are formed in a mirror: namely, that objects whose images are formed beyond the eye and the mirror appear behind the mirror, while those whose images are formed behind the eye appear in front of the mirror.²⁶

Here we have a conclusion that is both counterfactual and illogical by modern lights yet is forced upon Ptolemy by his theoretical model. When the intersection-point of the incident ray and the cathetus is indeterminate or lies at the center of sight, nothing is in fact seen. Nonetheless, while recognizing that nothing *appears* to be seen in these cases, Ptolemy is compelled to posit some sort of image. Why? Because at the very heart of his visual theory lies the supposition that, whenever a visual ray touches a visible object, the object will be seen. In all the above cases, the reflected ray BTHZ touches the object-points, so those points must all somehow be seen. To account for the *appearance* of invisibility in certain cases, then, Ptolemy suggests a transposition of the image by the visual faculty toward the mirror's surface. Coalescing with that surface in cases 1 and 4, the image takes on its color and thus becomes indistinguishable from it. In short, it becomes a true ghost image. In case 5, however, the image is transposed not to the reflecting surface itself but to a point between that surface and the center of sight so that the image created by the visual faculty is "real." Ptolemy concludes in *Optics*, IV, 71-73 (pp. 194-195) with an empirical corroboration of his geometrical analysis:

[3.19] [EXPERIMENT IV.1] But we can examine the cases of image-location that we have discussed, each in turn, using the bronze template we described earlier.²⁷ As before, in figure 3.28 [redrawn as 3.29], we draw lines EB and ZB [reflected from ABG] at equal angles and DH parallel to EB . We then draw EDZ , position the eye at E , and place the concave mirror on arc ABG while setting ourselves to observe the images of small colored pegs placed upon line BZ .

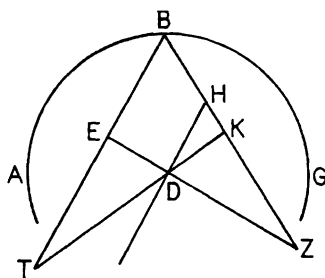


figure 3.29

The image of the peg placed at Z or H will seem to coalesce [with the surface] where point B lies on the mirror, and it will take on both its position and color. For there is no proper, determinate location for the image of anything located at Z or H . In fact, nothing lying at point H has a proper intersection-point, whereas what lies at Z shares a common location with the eye, since the intersection lies at E .

The images of objects lying between these two points [Z and H] do not appear on the mirror itself, because those images have determinate locations. But whatever lies between B and H appears to lie beyond the mirror at the point of intersection, whereas whatever lies between H and Z appears to lie in front of the mirror for the reason appropriate to it. For if we draw line BET , and lines DK and DZ as well, and if their intersections lie at points E and T , the image of point Z will lie at point B . Accordingly, the image of the intersection at point T will appear toward K , which lies in front of B with respect to D , because the location of T with respect to E lies in that direction.

This experimental demonstration is instructive, because it shows that Ptolemy got the “facts” straight at an observational level, even if theoretical considerations forced him to misconstrue them. Thus, in the case of an indeterminate intersection or one that occurs at the eye itself, no image is in fact seen, whereas if the intersection lies beyond the eye, one sees a sort of real image that is more or less distinct depending on how far beyond the eye the intersection lies. Bear in mind, however, that the fact that no images can be seen in certain circumstances was logically impossible for Ptolemy, so he was compelled to adduce the counter-fact of ghost images to account for the apparent lack of visible images.

4. Image-Formation and Distortion

The very formation of images in mirrors constitutes a distortion of sorts, insofar as the object appears to be in a place other than the one it actually

occupies. Such displacement constitutes a fundamental anomaly, insofar as there is a difference between apparent and real location. This fundamental anomaly is subject to specific variation according to the shape of reflecting surface, being least variable in plane mirrors and most variable in concave mirrors. Likewise, size- and shape-distortion constitute basic anomalies that vary with the shape of the reflecting surface, and so, to some extent, do the anomalies of image-reversal and image-inversion. Image-distortion thus ranges in complexity according to the progression of mirrors from plane, through convex, to concave. Our study in this section will follow that order of complexity, starting with plane mirrors, moving to convex mirrors, and concluding with concave mirrors.

Plane mirrors: the basic anomalies of image-formation in plane mirrors are analyzed in a somewhat cursory fashion in two propositions of the Euclidean *Catoptrics*. Proposition 7 (pp. 298-300), addresses image-inversion along the normal (i.e., in an up-down sense), whereas proposition 19 (pp. 314-16) deals with right-left reversal as well as with the size- and distance-relation between object and image:

[3.20] [PROPOSITION 7] *In plane mirrors altitudes and depths appear upside-down.*

Let AE [in figure 3.30] be an altitude, AL a plane mirror, and B the eye. Let visual rays BC and BD be reflected to E and K [respectively]. Hence, if the [incident] rays [BC and BD] are extended in a straight line, point E at the [object's] top appears at point T at the [image's] bottom, whereas point K, which is lower [on the object], appears at point Z, which is higher [on the image].

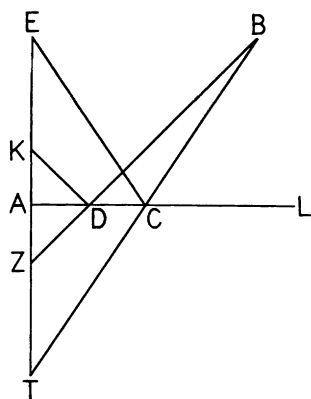


figure 3.30

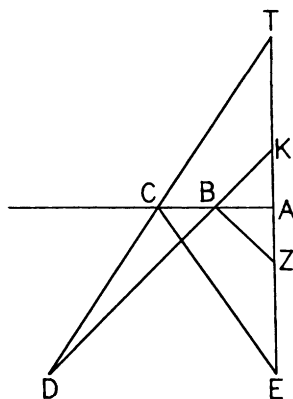


figure 3.30a

On the other hand, let EA [in figure 3.30a] be a depth, AC a plane mirror, and D the eye. Let rays DC and DB be reflected to E and Z [respec-

tively]. If, as before, the [incident] rays are extended [in a straight line] to T and K, point E, which is lower [on the object], will appear at T, which is higher [on the image], whereas Z, which is higher [on the object], will appear at K, which is lower [on the image].

[PROPOSITION 19] *In plane mirrors rightward objects appear leftward, whereas leftward objects appear rightward, the image is the same size as the visible object, and the distance [of the two] from the mirror is equal.*

Let AC [in figure 3.31] be a plane mirror and B the eye. Let visual rays BA and BC be reflected to E and D. Let ED be the object that is viewed, let

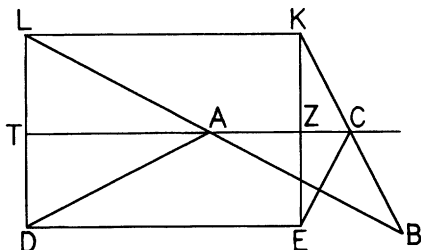


figure 3.31

perpendiculars EZ and DT be dropped from E and D to the mirror, and let them be extended [beyond]. Also, let rays BC and BA be extended and let them intersect the perpendiculars at K and L. Let LK be drawn. Therefore, E appears at K, while D appears at L, for this has already been demonstrated [in proposition 16, passage 3.16 above]. Hence, the leftward parts appear rightward, and the rightward leftward. Moreover, since angle KCZ = [angle] ZCE, and since the angles at Z are right by construction, ZK = ZE. Consequently, the distance of ED from the mirror is equal to the distance of image KL [from the mirror]. Finally, the object ED that is seen is the same size as its image KL, because EZ = ZK, DT = TL, and TZ is common and perpendicular.

From these two theorems we can conclude that, whereas the image's orientation is subject to distortion (i.e., left-to-right in the frontal plane and top-to-bottom in the plane of the normal), its relative distance and size are not. But what about shape? Ptolemy addresses this question in *Optics*, III, 82-86 (pp. 157-159):

[3.21] [THEOREM III.7] In plane mirrors, the image of a magnitude appears to have the same shape as that magnitude would if it were to lie where the visual rays would strike unreflected.

Let ABGD [in figure 3.32] be a right-angled parallelogram, let its sides be bisected at points E, Z, H, and T, and let the two lines EKH and ZKT be drawn between them. Let two circular arcs be drawn through point Z such that one of them, LZM, has its convex face toward the mirror and the re-

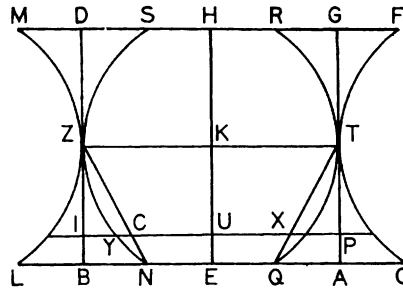


figure 3.32

flected ray, while the other, **NZS**, is identical but has its concave face toward them, and let them be positioned in such a way that **KZ** is perpendicular to them both, as is the case in a face-to-face orientation. Let line **EKH** lie on the plane mirror, and let the [reflected] visual ray proceed along **KZ** in the direction of the visible object.

Now let the two lines **AE** and **GH** be extended so that lines **EO** and **HF** are equal to lines **EL** and **HM** [respectively], and let lines **EQ** and **HR** be taken equal to lines **EN** and **HS** [respectively]. From what we have previously determined, then, it is evident that the image of line **BD** will appear as a straight line along **ATG**. Meanwhile, the image of convex arc **LZM** will appear along **FTO**, while the image of concave arc **SZN** will appear along **QTR**.²⁸

Line ATG is straight, because lines EA, KT, and GH are equal, and all the angles are right.

The line passing through points O, T, and F is convex, whereas the line passing through points Q, T, and R is concave. The reason is that the visual ray directed along line KZ to point T is normal to all the images of the lines, whereas the rays that reach the endpoints of the lines and all the remaining spots on them are oblique. With respect to ray [KZT] that reaches T, the rays that reach O and F are proportionately longer than those that reach A and G, whereas with respect to ray [KZT] that reaches T, the rays that reach Q and R are proportionately shorter than those that reach A and G. In the case of magnitudes or lines to which visual rays are dropped, when the ratio of oblique to orthogonal rays is greater [than it is for a straight line] from a face-to-face perspective, then those magnitudes will appear convex. When this same ratio is smaller [than it is for a straight line], those magnitudes will appear concave. Line ATG will therefore be convex when it is curved to coincide with OTF, whereas it will be concave when it is curved to coincide with QTR. So line ATG [itself] is straight, whereas without exception line OTF is convex and line QTR is concave.

Distortion with respect to image-location, finally, is treated by Ptolemy in *Optics*, III, 91-95 (pp. 159-161):

[3.22] [THEOREM III.8] In plane mirrors, the image lies on the same side as the actual object whose image it is, and if visible objects move in some given direction, their images move in the same direction as far as the eye is concerned.

Let straight line **ABG** [in figure 3.33] lie on a plane mirror, and let **D** be the eye and **EZ** the visible object. Let the two ray-couples **DAE** and **DBZ** [emanating] from the eye at **D** be reflected at equal angles to points **E** and **Z**. Then let **DA** and **DB** be continued rectilinearly to points **H** and **T** [respectively], where they intersect the two perpendiculars [**EH** and **ZT**] dropped to **AB** from points **E** and **Z**. Let line **HT** be joined. Point **E** will therefore be seen at point **H** and point **Z** at point **T**, and, in fact, those two images appear to lie on the same sides as the two points generating them.

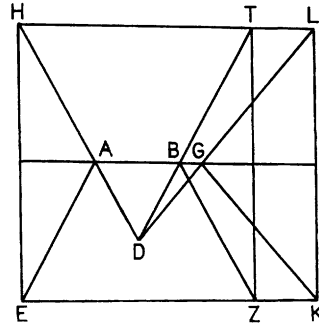


figure 3.33

Now, let line **EZ** be continued rectilinearly until **Z** reaches **K**, and let ray **DGK** [emanating] from point **D** reflect at equal angles to **K**. Then let line **DG** be extended until it meets point **L** on the perpendicular dropped from point **K** to line **AG**. The image that appeared at **T** has thus been shifted to **L**, and this shift takes the same direction as that of the visible object itself, i.e., from **Z** toward **K**.

For instance, if we assume that **KZ** and **BG** both lie above the eye, then image **TL** will appear to lie higher than the eye. And this image will be seen as the object itself is seen [in direct vision]: what is higher appears above, and what is lower appears below. Objects that are seen directly appear in precisely the same way: i.e., what is above appears above, and what is below appears below, for a higher visual ray sees what is above, whereas a lower one sees what is below.

If, moreover, we assume that both **KZ** and **BG** lie to the right of the eye, then image **TL** will also lie to the right, and **T** will shift to the right [when it moves] toward **L**. Nevertheless, the image of **T** will be judged to lie on the right-hand side [of **TL**], while that of **L** will be judged to lie on the left-hand side, for when objects are seen directly, face-on and parallel to our facial plane, their right side does not correspond to [our] right side, nor does their left side correspond to [our] left side; instead, the reverse is true.

In all, then, image-distortion in plane mirrors is as simple as it is limited: object and image are symmetrically related in terms of orientation and location, whereas relative size, shape and distance remain unaffected.

Convex mirrors: it is with convex mirrors that we encounter true image-distortion—distortion, that is, of shape, size, and distance rather than simply relative orientation and location, which are in fact “distorted” by image-reversal in the same way as they are in plane mirrors: i.e., left-to-right in the frontal plane and top-to-bottom in the plane of the normal. This top-to-bottom reversal in the plane of the normal is dealt with in extremely cursory fashion by Euclid, in proposition 8 of the *Catoptrics* (p. 300), where the burden of proof is laid upon proposition 7 (passage 3.20 above), for, since it has been established that the visual rays reflected to the top and bottom of a given altitude do not intersect (passage 3.4 above), “the rest follows as [it does] in plane mirrors.”

As to the left-right reversal of images in convex mirrors, Euclid addresses this issue in the first part of proposition 20 of the *Catoptrics* (pp. 316-318):

[3.23] [PROPOSITION 20] *In convex mirrors left appears right, whereas right appears left. . . .*

Let AC [in figure 3.34] be a convex mirror, T the center of the sphere, and B the eye. Let rays BA and BC be reflected to D and E, and let DE be the object that is seen. Then, from centerpoint T let TD and TE be drawn to D and E, let the rays [BA and BC] be extended to Z and H, and let image ZH be drawn. Accordingly, D appears at H, and E at Z. Therefore, the

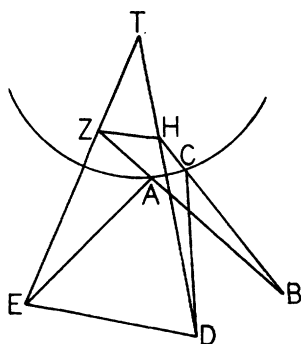


figure 3.34

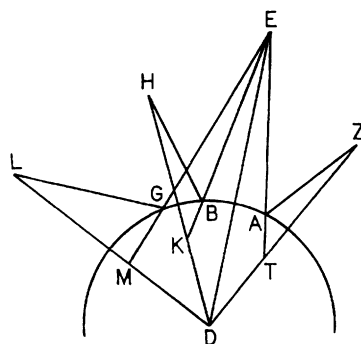


figure 3.35

right-hand side [of DE] appears [to lie] on the left [of the facing image], whereas the left-hand side appears on the right.

The problem of image-reversal along the frontal plane is treated by Ptolemy in a more comprehensive way, as is evident from the following theorem, which comes from *Optics*, III, 127-130 (pp. 170-171):

[3.24] [THEOREM III.16] *In convex mirrors, the image of objects is seen on the same side as the actual object, and, if the visible objects move in a*

given direction, their image moves in the same direction.

Let ABG [in figure 3.35] be the arc of a circle lying on a convex mirror whose center is D , let E be the eye, and let the visible object occupy points H and Z on both sides of the eye. Let the two ray-couples EAZ and EBH be reflected at equal angles to those points, let lines DZ and DH be drawn, and let EA and EB be extended to meet them at points T and K [respectively]. Accordingly, Z will be seen at T , and H at K , and each of those images [will lie] on the same side as the actual object of which it is an image.

Let H be moved toward L , and let ray-couple EGL be reflected to it at equal angles. Then let line DL be drawn so that line EG , when extended, intersects it at M . The image that lies at point K is thus transposed to M , so it is transposed to the same side as the object itself; and this follows necessarily from the reasons that we gave in our discussion of the image seen in plane mirrors.

If we place points H and L above the eye, then points K and M , which are the images of H and L , ought to appear above the eye, and the magnitude that lies between them will appear above the eye. And if we place H and L to the right of the eye, their images K and M will lie to our right, and so [the magnitude that lies between them] will be judged to lie to the left. Still, this image will not be seen in the same way as an object that faces us directly, but [what lies to the right] will be judged to lie to the left, and what lies to the left will be judged to be right according to what is usual, as we have already said, in the case of facing objects.

As with the problem of image-reversal in the frontal plane, so with the problem of distance-distortion, Euclid's treatment in the second part of proposition 20 of the *Catoptrics* is relatively primitive. Ptolemy, on the other hand, provides a fairly complete analysis of the problem in *Optics*, III, 110-116 (pp. 165-166):

[3.25] [THEOREM III.12] In convex mirrors, the distance of the visible object from the eye, as well as from the mirror's surface, is greater than the distance of the object's image [from the mirror's surface].

Let $ABGD$ [in figure 3.36] be the arc of a circle lying on a convex mirror, and let the center be at E , the eye at Z , and the visible object at H . Let the visual ray-couple ZBH be reflected at equal angles to the visible object, and let the two lines EB and EH be joined. Then let BZ be extended until it intersects EH at point T . Accordingly, the image of point H will lie at T . We say, therefore, that lines BZ and BH [taken together] are longer than line ZBT , and that GH is longer than GT .

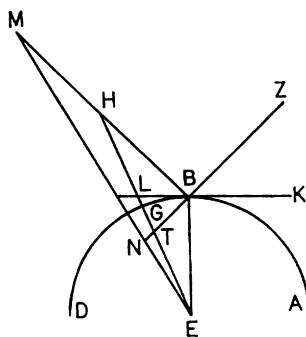


figure 3.36

Let line **KBL** be drawn tangent to the circle at point **B**. Since angle **ABZ** = angle **DBH**, and [horn] angle **ABK** = [horn] angle **DBL**, it follows that angle **KBZ**, which equals angle **LBT**, is equal to angle **LBH**. But angle **LBT** is acute, because angle **EBL** is right, and angle **BLH** > angle **BLT**. Therefore, line **BH** > line **BT**. With line **ZB** taken as common, lines **ZB** and **BH** [taken together] will be longer than line **ZBT**.

Moreover, since **BH** : **BT** = **LH** : **LT**,²⁹ line **LH** > line **LT**. **GH** therefore surpasses line **GT** by an even greater margin.

On the basis of what we have said, it will also be easily demonstrated that, in the case of objects that withdraw [away from the mirror's surface] or objects that lie farther than others [from that surface], the image lies farther away, and its distance [behind the mirror] appears greater.

Accordingly, let line **BH** be extended to **M**, and let line **EM** be joined. Then let line **ZBT** be extended until it meets line **EM** at point **N**.

If, then, we assume that **H** is distinct from **M** and that **M** is farther away [from the mirror's surface] than **H**, then **N**, which is the image of **M**, will certainly lie farther away from the eye at **Z** than **T**, which is the image of **H**. Moreover, if we assume that **H** coincides with **M** but that its distance increases as it moves from **H** to **M**, then its image will also become more remote, since it has shifted from **T** to **N**.

Next comes the issue of size-distortion, with which Euclid deals in proposition 21 of the *Catoptrics*, where his proof is indirect, depending, as it does, on showing that the image in a given convex mirror will be smaller than the image in a corresponding plane mirror, this latter image being the same size as the object. As is evident in the following excerpt from *Optics*, III, 117-120 (pp. 166-167), Ptolemy approaches the problem of size-distortion in a more direct way:

[3.26] [THEOREM III.13] In convex mirrors, when objects are situated in the way we specified for plane mirrors (where the lines joining the endpoints of the object lie face-on to the mirror), the image appears smaller than the objects themselves [would appear] if they were transposed to where the image lies, with the very same orientation and distance, and were viewed without reflection.

Let **ABG** [in figure 3.37] be the arc of a circle on a convex mirror, whose center is **D**, and let **E** be the eye. Let normal **EBD** be dropped from point **E**, let **ZH** be the line joining the endpoints of the visible object, and let the visible object be so disposed that line **EBD** bisects line **ZH** at right angles, as is the case for a facing object. Let the two ray-couples **EAZ** and **EGH** [emanating] from point **E** be

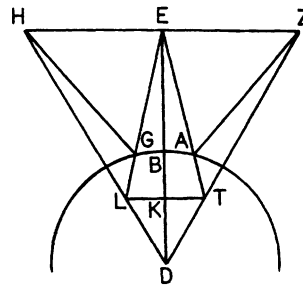


figure 3.37

reflected at equal angles to **Z** and **H** [respectively]. Then, let lines **DZ** and **DH** be drawn, and let **EG** and **EA** be extended until they meet lines **HD** and **ZD** at points **L** and **T** [respectively]. Finally, let line **TL** be joined. According to the principles we have established, then, the image of **Z** ought to be seen at **T**, and the image of **H** at **L**. Also, line **TL** will be the line joining the endpoints of the object's image, and its orientation, [being face-to-face,] will be the same as that of **ZH**.

Furthermore, since the distance of point **E**, the eye, from either of the points **Z** or **H** is the same because of their similar disposition, then the angles created by reflection to those points will be equal. Also, the distances of points **T** and **L**, which lie on the image that is seen in the mirror, will be equal with respect to point **E**, so that lines **ET** and **EL** will be equal. Meanwhile, angles **ETK** and **ELK** will be equal, as will angles **TEK** and **KEL**, so triangles **EKT** and **KEL** will be equal and equiangular. But the angles at **K** are right, so line **TL** is parallel to line **ZH**. Consequently, $DH : DL = ZH : TL$. But line $DH > DL$, so line $ZH > TL$.³⁰

Now, if line **ZH** were transposed to the location of line **TL**, and if the eye at **E** were to view it directly, it would be seen under a larger visual angle than **TEL**, which [image **LT**] subtends. And this accords with what we proposed.

The final issue to be dealt with is that of shape-distortion. As usual, Euclid's treatment (in proposition 23 of the *Catoptrics*, pp. 324-326) is both sketchy and incomplete, covering only the basic distortion of plane figures by convex mirrors. Ptolemy, on the other hand, takes a more comprehensive approach in *Optics*, III, 121-126 (pp. 167-170):

[3.27] [THEOREM III.14] In convex mirrors, facing straight lines appear convex. In the case of circular arcs, however, those whose convex curvature faces the mirror and the reflected ray appear convex, whereas those whose concave curvature faces the mirror sometimes appear convex, sometimes straight, and sometimes concave.³¹

Let **ABG** [in figure 3.38] be the arc of a circle lying on a convex mirror, whose center is **D**, and let **E** be the eye. Let normal **EBD** be dropped from the eye to the mirror, and let two equal arcs **BA** and **BG** be marked off on either side of **B**. Let the two ray-couples **EAZ** and **EGH** [emanating] from the eye at **E** be reflected at equal angles from points **A** and **G** [re-

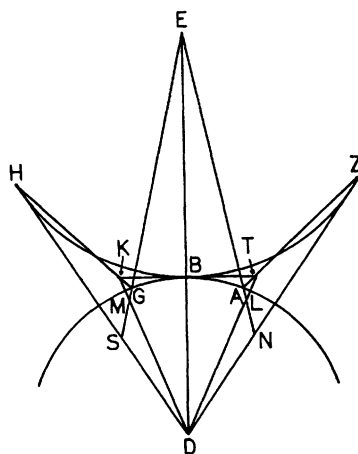


figure 3.38

spectively]. Let two lines be drawn tangent to the circle at point **B**, and let one of them, **TBK**, be straight, while the other, **ZBH**, is a circular arc with its convex curvature facing the mirror and the reflected ray. Then, let lines **TD**, **KD**, **ZD**, and **HD** be drawn to intersect [the extensions of] lines **EA** and **EG** at points **L**, **M**, **N**, and **S** [respectively]. Accordingly, the image of point **T**, which is one endpoint of the straight line, will lie at **L**, and the image of point **K** [the other endpoint] will lie at **M**. Meantime, the image of **Z**, which is one endpoint of the convex line, will lie at **N**, whereas the image of **H** [the other endpoint], will lie at **S**.

It is therefore evident that for both [line-segments] common point **B** will appear at **B**, since this is where the mirror's surface and the visual ray intersect. However, the convexity of arc **ABG** lies toward the eye at **E**, and the images of simple lines, which comprise a single line, themselves comprise a single line of a unique type.³² Hence, the eye will see the two lines passing through points **L**, **B**, and **M** and points **S**, **B**, and **N** as convex, since they are more sharply curved than arc **ABG**. Still, arc **SBN** will be more sharply curved than **LBM**. Since the objects face the eye, and since the oblique lines produced from [the endpoints of] the diameter are proportionately longer with respect to the normal [in the case of the image] than the corresponding oblique rays when the object is seen directly, the line passing through points **L**, **B**, and **M**, which defines the image of straight line **TBK**, will be convex. By the same token, the line passing through points **N**, **B**, and **S**, which defines the image of convex line **ZBH**, will also be convex.

[THEOREM III.15] If the line that is seen is concave, it can be demonstrated as follows how it is possible, using the distance along the normal between the two arcs [of the concave object-segment and the convex mirror], that the image might sometimes appear convex, sometimes concave, and sometimes straight:

Let **ABG** [in figure 3.39] be the arc of a circle lying on a convex mirror whose center is **D**, let the eye be at **E**, and let the normal be **EBD**. Let circular arc **AZG** be drawn from points **A** and **G** so that it cuts perpendicular **BE** [at **Z**]. Let two equal arcs **ZH** and **ZT** be marked off on each side of **Z**, and let the two ray-couples **EKH** and **ELT** [emanating] from point **E** be reflected at equal angles to **H** and **T**. Let **DT** and **DH** be drawn, and let them intersect the extensions of lines **EK** and **EL** at points **M** and **N**. Thus, **H** will be seen at **M**, and **T** at **N**, while points **A** and **G** will be seen where they actually lie.

Depending on how sharply curved concave arc **AZG** is, it is possible for points **M** and **N** sometimes to lie between arc **ABG** (which is an arc on the mirror) and the

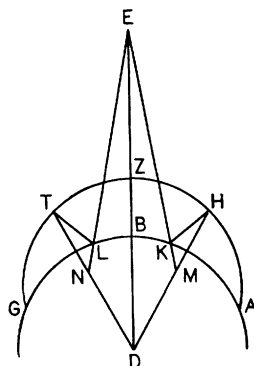


figure 3.39

straight line joining points A and G, sometimes to lie on that straight line itself, and sometimes to lie between that straight line and D (which is the sphere's center). Now it is clear that, when images of this kind lie on one line, and points M and N lie between the mirror and the straight line joining A and G, then the eye at E will see their convexity directed toward it. But if points M and N lie on that same straight line [joining A and G], then the image will appear straight, whereas if the image is farther removed from E [than that straight line], it will appear concave.³³ This will be obvious through a construction of the image on the basis of the ratios between oblique and normal rays.

Concave mirrors: Not surprisingly, the variability of image-location in concave mirrors, which depends upon the relative locations of the eye, the object, and the mirror's surface, makes image-distortion correspondingly variable and complex. As in the case of spherical convex mirrors, so in this one, apparent distance, apparent size, and apparent shape are all subject to distortion. However, in the case of spherical concave mirrors, the type of distortion depends, first, upon whether the image is "virtual" (i.e., lies behind the mirror's surface) or "real" (i.e., lies between the eye and the mirror's surface). Also, depending on the type of image, there can be image-inversion.

Euclid deals with the issues of distance- and size-distortion, as well as image-inversion, in the second part of proposition 28 of the *Catoptrics* (pp. 334-336):

[3.28] Let the eyes B and C [in figure 3.40] be placed above the point [N] that bisects the radius [ZA], and let them be placed equidistant from radius ZA. I say that B and C are seen [in the mirror], that the right-hand side appears leftward, while the left-hand side appears rightward, that the image of the face appears larger [than it should], and that the image appears to lie at a greater distance from the mirror[’s surface] than it should.

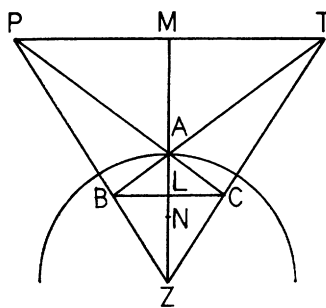


figure 3.40

Now, let BA be a reflected ray, let ZB and ZC be extended, and let BA be extended. With AZ having now been bisected at N, $BZ > BA$, and angle $BAZ > BZA$. But angle $BAZ = \text{angle } CAZ$, so angle $CAZ > \text{angle } BZA$. Hence, if ZB and CA are extended, they will intersect.³⁴ Let them intersect at P. By the same token, BA and BC will intersect at T. Therefore, C will be seen at T, and B at P, and what lies on the right appears to the left, and what lies on the left appears to the right. Furthermore, $TP > BC$, for they are parallel. Therefore, the image appears larger and more distant from the

mirror[s surface than it should], for $MA > AL$.

If, however, the eyes are placed beyond the [mirror's] diameter, then the right-hand side appears rightward, whereas the left-hand side appears leftward, and the image of the face appears smaller [than it should] and [is seen] between the face and the mirror[’s surface].

Let B and C [in figure 3.41] be the eyes, Z the mirror's centerpoint, and let AZD be perpendicular to the diameter as well as to BC so that $BA = AC$. Let [visual ray] BD be reflected to C , let BZK and CZE be extended through the [mirror's] center [of curvature Z], and let KE connect K and E . Hence, B is seen at K , and C at E . Accordingly, the right-hand side appears to the right, the left-hand side to the left, image EK appears smaller than the face BC , and the image appears between the face and the mirror[s surface].³⁵

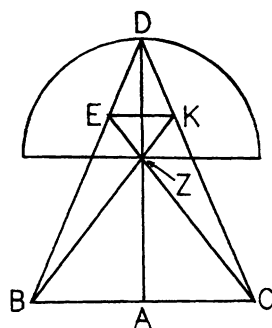


figure 3.41

As we might by now expect, Ptolemy's treatment of image-distortion in concave mirrors is far more systematic and extensive than Euclid's. Accordingly, he deals with each type of distortion in its turn, beginning with an analysis of distance-variation for point-images. The following passage from *Optics* IV, 109-119 (pp. 205-207) deals with two cases: the first involving images that appear behind the mirror's surface (i.e., "virtual" images), the second involving images that appear between the mirror's surface and the eye (i.e., "real" images):

[3.29] [THEOREM IV.30] Accordingly, in concave mirrors, when the image lies behind the mirror, the distance of the visible object [along the reflected ray-couple] will be smaller than that of its image [along the incident ray] if the visual radiation were to continue behind the mirror.

Let ABG [in figure 3.42] be the arc of a circle inscribed in a concave mirror with centerpoint E . Let point Z be the eye, and point H the visible object to which ray-couple ZBH is reflected at equal angles. Let [cathetus] $EHGT$ be drawn, and let BZ and HE be extended to intersect at point T beyond the mirror. Therefore, the image of H will lie at point T . We say, then, that $BZ + BH < ZT$,

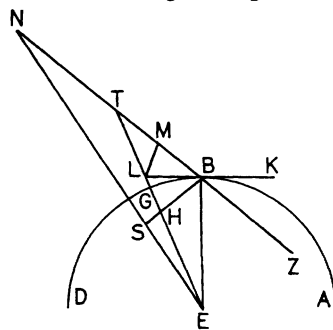


figure 3.42

whereas $GH < GT$.

Let [normal] EB be drawn, and let line KBL be drawn tangent to the circle at point B . Therefore, [curvilinear] angle ABZ = [curvilinear] angle GBH , and [horn] angle ABK = [horn] angle GBL , so that the whole angle KBZ , which is equal to angle TBL , is equal to angle LBH . Now angle BLH is acute, since angle LBE is right. Therefore, angle $BLH < angle BLT$. And if we posit angle $BLM = angle BLH$, then, since angles MBL and LBH in triangles MBL and LBH are equal, while side BL is common, it follows that $MB = BH$. Consequently, $BH < BT$. And if we take BZ as common, then $ZB + BH < ZBT$.

Moreover, since $LH : LT = BH : BT$, while $BH < BT$, then $LH < LT$, and GH will be much smaller than GT .

From what we have established, it is evident that, if the distance between the objects and the same viewpoint increases, or if that distance is greater, then the distance between the images and the eye increases, or that distance will be greater. For if we extend BH to S and continue [cathetus] ES until it meets the prolongation of ZBT at point N , then the image of S will lie at point N , and so $BS + BN > BH + BT$.

[THEOREM IV.31] If the image that appears in concave mirrors lies between the eye and the mirror, the distance of the visible object from the eye [along the reflected ray-couple] will be greater than the distance of its image [along the incident ray]; nevertheless, the image's distance from the mirror will sometimes be smaller, sometimes greater, and sometimes equal.

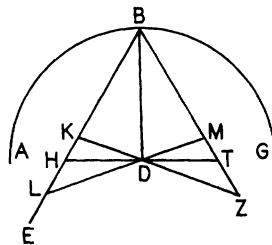


figure 3.43

Let ABG [in figure 3.43] be the arc of a circle inscribed in a concave mirror with centerpoint D , and let normal DB be drawn.

Let point E be the eye, and let ray-couple EBZ emanating from point E be reflected at equal angles. Then, let [cathetus] TDH be drawn through point D perpendicular to BD , and let the two catheti KDZ and LDM intersect one another at corresponding angles.

If, therefore, we place [magnitude] ZTM at the location that is observed, then, according to principles established earlier, the image of Z will lie at point K , the image of T at H , and the image of M at point L . Indeed, they fall between E and B , and their distance from E will be less than EB . Moreover, that distance will be much less than the [eye's] distance from the object-points reached by reflection, e.g., [along] EB and BM .

It is also evident that, if angle EBZ is bisected by line BD , then $BT = BH$, and $DT = DH$. But line $BZ > BK$, while $ZD > DK$. Also, line $BM < line BL$, and $MD < LD$. Hence, the image of T will lie at point H , and the distance of both [points] from point B will be equal. But the distance[from

point **B**] of **Z**'s image, which appears at **K**, will be less [than that of **Z** itself], and the distance [from **B**] of **M**'s image, which appears at **L**, will be greater [than that of **M** itself].

So it is clear that, when the distance of objects from the same eye increases, or when it is greater, the distance of their images from the eye increases or will be greater. For, since the distance of point **Z** from the eye is greater than that of the remaining points, its image appears farther away, and since the point-object at **M** lies nearer the eye, its image, which is at **L**, is formed nearer [the eye] than the remaining images.

On this basis, then, it is evident that, when the distance of the visible object is no greater than the distance of the line [**HDT**] passing through the center of the sphere, but when, instead, the object itself and the mirror lie on the same side of this line, the actual object will always lie a shorter distance from the mirror than its image.

Ptolemy then turns from point-images to lengths in order to establish a link between eye-to-image distance and size-distortion, that link following from the fact that, if the image appears farther away than it should, it will also appear smaller than it should. The following passage from *Optics*. IV, 120-129 (pp. 207-210) addresses this size-distance linkage in the case, first, of "virtual" images, and second, of "real" images:

[3.30] [THEOREM IV.32] In concave mirrors, if the image of a visible object appears behind the mirror and the location of the generating object is what we have called "facing," then, according to what we have already said in regard to other mirrors, the lines joining the endpoints of the images of visible magnitudes appear longer than the lines joining the endpoints of the visible magnitudes themselves, assuming that those magnitudes were observed directly from the [same] fixed viewpoint at the same distance as that of the image and with the same orientation.

Let **ABG** [in figure 3.44] be the arc of a circle inscribed in a concave mirror with centerpoint **D**. Let **E** be the eye, and let normal **DEBL** be drawn from **E**. Let **ZH** be the line joining the endpoints of the visible object, which is oriented so that line **BD** bisects it at right angles, for it must be placed in a directly facing position. Then, let the two rays **EA** and **EG** be reflected from **E** to **Z** and **H**. Let those two lines be extended to intersect the continuations of catheti **ZD** and **DH** at points **T** and **K**, which lie behind the mirror. Let line **KT** be joined, and let line **DB** be extended to cut it at point **L**. Therefore, the image of **Z** will be seen at **T**, and the image of **H** at **K**. But line **TK** joins the endpoints of the object's image, and its orientation is the same as that of

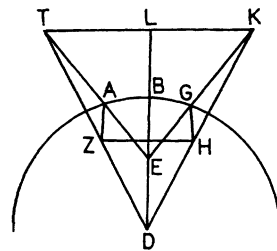


figure 3.44

ZH.

Therefore, since points *Z* and *H* lie the same distance from the eye at *E*, and since the angles of reflection are equal, image-points *K* and *T* will lie the same distance from point *E*. Also, since both lines *EL* and *ET* are equal to lines *EL* and *KE*, respectively, and since angle *TEL* = angle *KEL*, it necessarily follows that the [respective] angles of both triangles are equal. The angles at *L* will be right, and *KT* will be parallel to line *ZH*. Moreover, $KT : ZH = TD : ZD$. But $TD > ZD$. Therefore, $TK > ZH$. And if *ZH* is transposed to location *KT* and maintains the same orientation, visual angle *KET* will be larger than the visual angle subtended by *ZH*, assuming that it is looked at directly and that its distance and orientation are the same as that of *TK*. Thus, to the eye at *E*, *KT* will appear longer than *ZH*.

[THEOREM IV.33] In concave mirrors, when the image of a visible object appears between the mirrors and the eye, and when it is oriented [in a facing position] as we specified, the straight lines joining the endpoints of the visible magnitudes sometimes appear to be equal to [those of] their images, sometimes longer, and sometimes shorter, when those visible magnitudes are observed directly and maintain the same disposition and distance as their images, assuming that the eye remains fixed.

Let *ABG* [in figure 3.45] be the arc of a circle inscribed in a concave mirror with centerpoint *D*, let line *BD* be drawn, and let *E* be the eye. Let line *ZH*, which joins the endpoints of the visible object, be positioned in such a way as to be bisected by line *DB* at right angles, according to what must obtain in the case of things that have a facing orientation. Let the two ray-couples *EAH* and *EGZ* emanating from point *E* be reflected at equal angles to points *Z* and *H*. Let [catheti] *Z[D]T* and *K[D]H* be drawn, and let them cut lines *EA* and *EG* in points *T* and *K*. Finally, let line *KLT* be joined. The image of *Z* will therefore appear at point *T*, and the image of *H* at point *K*. *KT* will be the line joining the endpoints of the image, and its orientation will be the same as that of *ZH*.

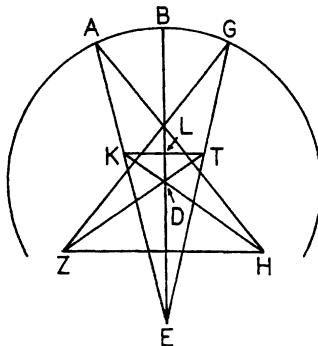


figure 3.45

Therefore, since the distances of the eye at *E* from points *Z* and *H* must be equal, as we have previously said, and since the angles of reflection are equal, the distances of *K* and *T* from point *E* are equal. Moreover, the two lines *KE* and *EL* are equal to lines *ET* and *EL*, and angle *KEL* = angle *TEL*. Therefore, triangles *ETL* and *EKL* will be equiangular, and the angles at *L* will be right, while line *TK* will be parallel to line *ZH*. Also, $ZH : KT = DZ : TD$. But we have already shown in the previous discussion that it is pos-

sible for DZ sometimes to be equal to DT , sometimes shorter, and sometimes longer. Thus, it is possible for ZH sometimes to be equal to KT , sometimes longer, and sometimes shorter.

However, if ZH is transposed to the position of KT , while it maintains the same orientation, angle KET will sometimes be equal to the angle subtended by ZH , if its disposition and distance are the same as that of KT and if it is observed directly, and it will appear equal to the eye at EL . But sometimes it will be greater and will appear greater, whereas sometimes it will be smaller and will appear smaller.

Moreover, all objects whose distances from point E are less than half the radius of the sphere that forms the mirror maintain a distance between the endpoints of the magnitude that is greater than the distance between the endpoints of the image.

Now we come to the issue of shape-distortion, which Ptolemy addresses for both cases in order of simplicity (i.e., first for “virtual” images and then for “real” images) in *Optics*, IV, 130-141 (pp. 210-214):

[3.31] [THEOREM IV.34] When the image of an object that is seen in concave mirrors lies behind them, or between them and the eye, straight objects that face them directly appear concave. Furthermore, circular objects whose concave side faces the mirror and the reflected ray appear concave, but those whose convex side faces the mirror sometimes appear concave, sometimes straight, and sometimes convex.

[CASE 1]

First, then, let the image of the visible object lie behind the mirror. Let ABG [in figure 3.46] be the arc of a circle inscribed in a concave mirror with centerpoint D , and let line BDE be drawn. Let line ABG be bisected at point B . Then, through points A and G , let straight line AZG be drawn along with circular arc AHG , which has its concave side facing the mirror and the reflected ray. Let $DH > HB$, and let the eye be at point E .

The image of points A and G on these two lines will thus appear at points A and G themselves. Meantime, the image of points Z and H will appear behind the mirror, Z at point K and H at point L , for it has been demonstrated that the farther objects lie from the same viewpoint, the farther their images will lie from that viewpoint. But the distance of H from the eye at E , as judged along the ray-couple $[EB, BH]$ reflected from point B , is greater than the distance of Z from the eye at E [along ray-couple EB, BZ]. Point L is also farther away than point K , and both of those points lie on the image.

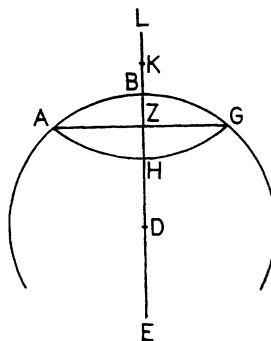


figure 3.46

So the image of line *AZG* will appear on the line that passes through points *A*, *K*, and *G*, whereas the image of arc *AHG* will appear on the line passing through points *A*, *L*, and *G*. Moreover, the concavity of arc *ABG* will appear to be turned toward center of sight *E*, and the lines passing through points *AKG* and *ALG* will be more pronounced in their concavity than *ABG*, since the rays falling perpendicular to those arcs are proportionately longer than the ones falling more obliquely. Generally speaking, though, the images of straight and concave objects will appear concave.

To continue, let a convex arc *EBZ* [in figure 3.47] be drawn through point *B*, tangent to the circle and [thus] to arc *ABG*. Let [normal] lines *DEH* and *DZK* be drawn, and let *DE* and *DZ* be longer than *EA* and *ZG*. Let tangent *LBM* be drawn through point *B* of the circle.

Accordingly, the image of *EZ* can sometimes be seen between *L* and *A* and between *M* and *G*, sometimes on *L* and *M* themselves, and sometimes beyond points *L* and *M* (e.g., at points *H* and *K*), all depending on the distance of *E* from *A* and of *Z* from point *G*. But the image of *B* will lie at point *B* itself.³⁶

If the images of *E* and *Z* lie at points *L* and *M*, then the image of convex [arc] *EBZ* will appear along straight line *LBM*. But when the images of *E* and *Z* lie between *A* and *L* and between *G* and *M*, then the image will appear concave, for when oblique rays falling on objects are proportionately smaller than perpendicular ones, those objects appear concave. Finally, if the image[s] of *E* [and] *Z* lie beyond points *L* and *M*, then the image will appear along a convex line, for in this case the opposite to what we just specified happens insofar as the oblique rays are proportionately longer than the perpendicular ones. Thus, when something is concave in comparison to straight line *LBM*, it will be absolutely concave, and when it is convex in comparison to that [same straight line], it will be absolutely convex.

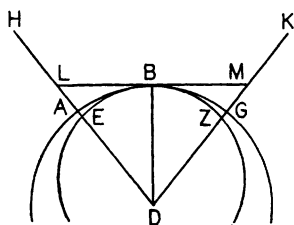


figure 3.47

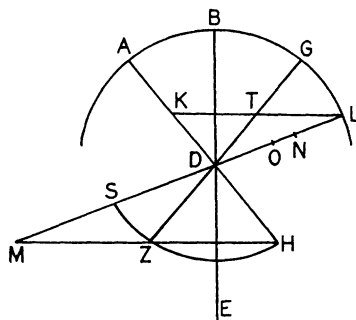


figure 3.48

[CASE 2]

Now, let the image lie between the eye and the mirror. Let *ABG* [in figure 3.48] be the arc of a circle inscribed in a concave mirror with centerpoint *D*, let line *EDB* be drawn normal to the mirror, and let the eye be at point *E*.

Let HZ be the line joining the endpoints of the visible object, and let line ED bisect it at right angles.

Accordingly, when catheti ZDG and KDH are extended, the image of Z will appear at point T , and the image of H at point K . Let cathetus TK be drawn, and let it be extended to point L on the arc. Let line LDM be drawn, and let it intersect the extension of line ZH at point M . Thus, the image of point M will appear between points D and L , since none of the rays emanating from the eye is reflected at equal angles from point L to point M .³⁷ But if we suppose that the image of point M lies at point N , then the image of straight line MH will be seen along the [curved] line passing through points K , T , and N , and the concave side of this image will face the eye.

The same thing holds if we assume that the line passing through points Z and H is circular, with its concave side facing the mirror. For, if we construct the arc to which the visual ray is reflected and have its concave side facing the mirror, as is represented by arc HZS , then the image of point S will lie between points D and N (e.g., at point O). And the line passing through points K , T , and O will be even more concave [than the previous one through K , T , and N]. But this very line $[KTO]$ lies on the visible object's image.

Let us now suppose that arc ZSH [in figure 3.49] has its convex side facing the mirror, let us draw straight line ZH joining the endpoints of the visible object, and let KNT represent its concave image. When the rays are shorter, the image of the object seen by them lies closer [to the eye]. Therefore, since the distance of S from the eye at E is less than that of L [from E], as measured by the ray-couple[s] EB , BS and EB , BL] reflected from point B , and since the distance $[EM]$ of its image, which lies at M , is less than that of L [along EN], then point S will be seen between D and N .

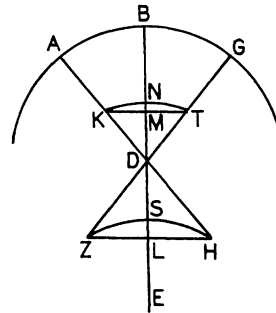


figure 3.49

So the image of S can sometimes appear at M , and the image of ZSH will lie on straight line TMK . But sometimes [it can appear] between M and N , and the line will appear concave. Finally, the image can sometimes appear between points D and M , and the line will appear convex, depending on how pronounced the concavity [of the visible object] is. Still, the image of [end]points Z and H will remain fixed at points T and K .

The final type of image-distortion involves image-inversion, which does not apply to “virtual” images but does apply to “real” ones under certain circumstances. Ptolemy’s analysis is found in *Optics*, IV, 142-151 (pp. 214-216):

[3.32] [THEOREM IV.35] In concave mirrors, when the image of a vis-

ible object lies behind the mirror, it is perceived to be on the same side as the actual object; and when visible objects are moved in any direction, their images appear to move in the same direction.

Accordingly, let ABG [in figure 3.50] be the arc of a circle inscribed in a concave mirror with centerpoint D . Let the eye be at E , and let Z and H represent two visible objects flanking it. Let catheti DZT and DHK be drawn, and let the two ray-couples EAZ and EBH be reflected at equal angles from E to Z and from E to H . Then, let lines EA and EB be extended so as to intersect lines DT and DK at points T and K . Therefore, the image of Z will lie at point T , and the image of H at point K . Moreover, those images will be perceived to lie on the same side as the actual objects.

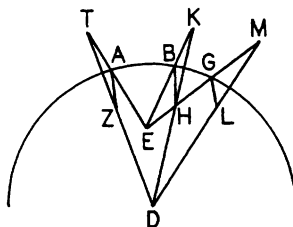


figure 3.50

Let H be moved to L , and let [ray-couple] EGL be reflected to it at equal angles. Then, let [cathetus] DLM be joined, and let it intersect the extension of EG at point M . Therefore, the image at K will be moved to point M , which lies in the same direction as that toward which the actual object has been moved.

Again, if H and L , which represent the visible objects, are above eye-level, then their images K and M will be above eye-level and will appear above [the original location]. For the upper parts of objects directly facing the viewer are apprehended by higher rays.

If, however, we suppose H and L to lie to the right of the eye, then their images, which are at K and M , happen to lie to our right. But those images are not judged to lie to the right, because the part of what we see facing us that is touched by right-hand rays and that appears to the right is actually on the left-hand side of what we see. And [so] the left-hand side of a facing visible object is apprehended by the right-hand rays in direct vision. But in the case of reflection, it is the right-hand side that is apprehended, and the facing sides of the images are directly opposite us, so the images that lie to the right are judged to lie to the left according to the normal arrangement of the visual rays in regard to position.

[THEOREM IV.36] In concave mirrors, when the image of a visible object lies between the mirror and the viewer, it is perceived as the reverse of the actual object, and when the actual object is moved in a given direction, its image appears to move in the opposite direction.

Let ABG [in figure 3.51] be the arc of a circle inscribed in a concave mirror with centerpoint D . Let the eye be at E , and let Z and H represent two visible objects on either side of the eye. From point E let the two ray-

couples EAZ and EBH be reflected at equal angles to those [two point-objects]. Let catheti ZDT and HDK be drawn, and let lines EA and EB intersect them at points T and K . Thus, the image of point Z will lie at point T , and the image of point H at K . Consequently, the images will appear on the opposite sides of their actual [generating] objects.

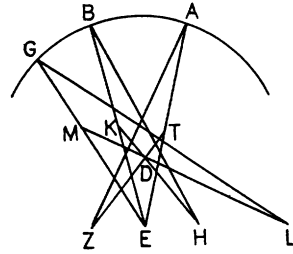


figure 3.51

By the same token, if point H is moved to point L , if ray-couple EGL is reflected to it, and if cathetus LDM is extended to meet line EG at point M , then the image of H , which is at K , will move in the opposite direction from that in which the actual object has been moved.

Furthermore, if H and L lie above eye-level, their images, which are represented by K and M , will lie below us, and the upper parts will appear lower. For things that are apprehended by lower visual rays appear at the lower side of the visible object, but what is apprehended in such reflections by lower rays appears at the upper side.

And if we suppose that H and L lie to the right of the eye, then their images (i.e., K and M) will appear to our left. And so right-hand objects will appear to the right, for the parts of facing objects that are apprehended by right-hand rays appear to the left, while those that are apprehended by left-hand rays appear to the right. Among objects seen by means of such reflections, moreover, those that move toward the right appear to move toward the left, because [in this case] our hand does not seem to move as a mirror-image, but in the opposite direction, as if it were a right hand [rather than a left]. Indeed, facing objects that are seen in direct vision are disposed so that their right sides lie to our left.

5. Visual Effects from Composite Mirrors

That there was a fascination among early optical theorists with mirror-effects, such as the repetition of images in composite mirrors or the unexpected appearance of images in certain mirror-arrangements, is clear from both Euclid's *Catoptrics* and, to some extent, Ptolemy's *Optics*. The latter, for instance, has three theorems (IV, 39-41) expressly devoted to image-repetition, and in fact so does the former (props. 13-15). The following theorem, prop. 13 of Euclid's *Catoptrics* (pp. 306-308), which is repeated (with elaboration) in Ptolemy's *Optics*, IV, 175-177 (pp. 224-225), serves as a canonical example:

[3.33] [PROPOSITION 13] *The same object can be seen in several mirrors.*

Let A [in figure 3.52] be the object that is supposed to be seen, let B be the eye, and let CD , DE , and EZ be three [plane] mirrors. Let perpendicu-

which the observer looks—and the other one, behind him, is moved toward and away [from his head], his visual ray will [eventually] reach the point where [the image of] his heel lies, and he will perceive himself as flying.

In the very next proposition, (prop. 16, pp. 352-356) Hero describes how to set up a mirror inside a room with a window so that what is going on outside in the street can be seen from within. This he follows with a description of a five-fold mirror arrangement in which someone standing in the middle will see himself in all five. He then turns his attention to the following problem: to set up a mirror so that, in approaching it, one sees not his own image, as he might expect, but another image entirely, one that seems to loom up at him. This effect is described in the following passage from chapter 18 of the *Catoptrics* (pp. 358-360):

[3.35] [CHAPTER 18] To set up a mirror in such a way that anyone approaching it sees neither himself nor anything else except some particular image that someone has preselected.

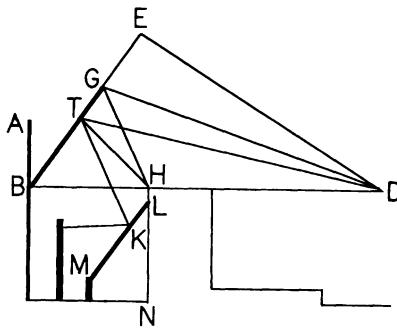


figure 3.54

Let **AB** [in figure 3.54] be a wall against which the mirror is to be placed, and let the mirror be inclined against it at some angle. At any rate, it should be sufficient if the angle is a third of a right angle. Then let **BG** be the surface of the mirror, and from **AB** let **BD** be drawn at right angles and extended to point **D**, where the center of sight is situated [and let **D** be placed] so that the perpendicular extended from it to mirror **BG** intersects [the extension of that mirror's surface] beyond it [i.e., at point **E**]. Let that perpendicular be **ED**, and to endpoint **G** of the mirror let **DG** be drawn, and let angle **BGH** = angle **EGD**. Hence, if a given ray from center of sight **D** touches endpoint **G** of the mirror, it will be reflected to **H**. Accordingly, from **H** let **HN** be drawn perpendicular to **DB**. Then let another ray, **DT**, strike the mirror, and let **TH** be drawn. Consequently, angle **BTH** > angle **ETD**. Thus, angle **GTD** = angle **BTK**, so **TK** intersects **HN**, and likewise all the rays that strike the mirror and are reflected from it intersect **HN**.

Therefore, let a plane mirror **LM** be placed inside **HN** and parallel to mirror **GB** so as to be struck by a ray reflected [from **GB**]. In that case, it is obvious that the eye will see nothing other than whatever lies inside **HN**, because all of the rays reflected [from **BG**] fall within **HN**. Therefore, let us place whatever image we like in front of plane [mirror] **LM**, and it is certain that none [of the images] of those approaching will appear, only the aforesaid image.

Despite first appearances, such an arrangement of mirrors to produce unexpected images had a practical purpose: as Hero makes clear in the remainder of the proposition, it was intended for use in temples where the effect would presumably arouse a sense of wonder among the mass of celebrants who could not fathom its underlying causes. Here, in short, we see technology in the service of religion (for an equivalent case involving burning mirrors see passage 5.7 below).

NOTES

¹“Definition” 3 of the *Catoptrics* says that “when a given perpendicular magnitude is viewed in a plane mirror, the straight distance between the mirror and the viewer is to the straight distance between the mirror and the perpendicular magnitude as the altitude of the viewer is to the altitude of the perpendicular magnitude.

²Presumably, this diopter would consist of either a narrow sighting-tube or a narrow sighting-slit through which the line of sight could be directed as narrowly as possible.

³This experiment could be carried out using all three mirrors at once if: 1) the concave mirror were shorter than the plane mirror which, in turn, were shorter than the convex mirror; 2) the axial pin were common to all three mirrors; and 3), the colored marker slid along arc **BE** were tall enough to be seen in all three mirrors at the same time. Otherwise, the experiment would have to be carried out for each mirror in turn.

⁴The point is that, if they did intersect, then an object placed at the intersection-point would necessarily be seen at two places behind the mirror along the extensions of incident rays **BC** and **BA**.

⁵Since angle **ORZ** + angle **R** = two right angles, and since angle **O** < angle **R**, then angle **O** + angle **ORZ** < two right angles, so lines **CZ** and **AI** must converge toward **I** and **Z** (by Euclid, *Elements* I, postulate 5).

⁶Since arc **ACT** > arc **CTK**, then angle **M** > respective angle **X**.

⁷This is the case treated by Hero, in chapter 9 of his *Catoptrics*.

⁸Euclid provides a specific complement to this proposition later on in proposition 24 of the *Catoptrics* (p. 326), where he undertakes to demonstrate that “in the case of concave mirrors, if the eye is placed at the center [of curvature], only the eye, by itself, appears.”

⁹Proposition 27 of the Euclidean *Catoptrics* makes the same basic point as this theorem.

¹⁰These are proportional sides of similar triangles **KLH** and **ENH**. The same applies to the next proportionality.

¹¹Since **KH** : **EH** = **KZ** : **ZE** and **KH** : **EH** = **KL** : **EN**, then **KZ** : **ZE** = **KL** : **EN**. But **KZ** : **EZ** = **KL** : **EM**. Thus, **KL** : **EN** = **KL** : **EM**, whence it follows that **EM** = **EN**.

¹²By Euclid, *Elements*, III, 7: “If on the diameter of a circle a point is taken which is not the center of the circle, and from the point straight lines fall upon the circle, that will be greatest on which the centre is, the remainder of the same diameter will be least, and of the rest the nearer to the straight line through the centre is always greater than the more remote . . .” For the full theorem, see *The Thirteen Books of Euclid’s Elements*, second edition, trans. Sir Thomas Little Heath, vol. 2 (1925; New York: Dover, 1956), pp. 14-15.

¹³By Euclid, *Elements*, III, 21: "In a circle the angles in the same segment are equal to one another," so, by extension, the angles in equal segments will be equal. For the full theorem, see *Elements*, trans. Heath, vol. 2, p. 49.

¹⁴The logic here is that, for all of the possible reflecting ray-couples from ZLH to ZAH [in figure 3.16a below] or from HMZ to HGZ, centerpoint D will not be included within the angle formed by those ray-couples. It follows that the normal from centerpoint D to the point of reflection will invariably fall outside that angle, so the angles of incidence and reflection (e.g., ZLA, HLM or HMG, ZML) cannot possibly be equal relative to their respective normal, DZL or DHM.

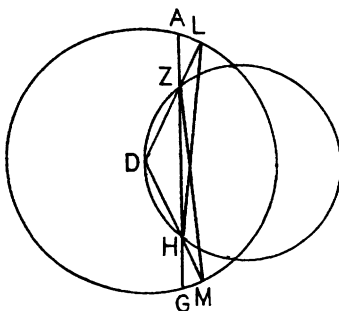


figure 3.16a

¹⁵By the previous theorem, no reflection can occur from point Z to point H from anywhere within arc AB nor from anywhere outside arc BK. Hence, by default, the reflection must occur from somewhere within arc BK.

¹⁶By Euclid, *Elements*, VI, 3.

¹⁷See note 14 above.

¹⁸If, for instance, we choose U as the point of reflection, then, given that angle ZUH will be bisected by normal UDC, it follows that $ZU : UH = ZC : CH$. Thus, ZU, which is shorter than ZN, will be longer than UH, which is longer than NH. By the same token, angle UCZ, which is subtended by the supposedly longer side UZ, should be greater than angle NLZ, which is subtended by the supposedly shorter side NZ.

¹⁹The point of this apparent excursus is to establish that, in spherical convex or concave mirrors, the image must lie on the perpendicular passing through the center of curvature from the object.

²⁰In other words, the image necessarily lies along the rectilinear extension of the incident ray.

²¹This is an empirical verification of the fact that the image lies on the cathetus of reflection.

²²This point is made explicitly in definition 4.

²³This is the first definition of the book. What it means, is that, when a line is straight, all its points lie in such a juxtaposition that, if the line were looked at endwise, the first point seen would block all the rest; cf. Euclid, *Elements* I, definition 4.

²⁴This point is made explicitly in definition 4.

²⁵Proposition 28 of Euclid's *Catoptrics* takes a different approach to the issue of image-determination. Euclid undertakes to prove there that, if the eyes are placed on either side of, and equidistant to, the radius of a concave mirror so that the line connecting them is perpendicular to that radius, and if a perpendicular is drawn to that radius through the point bisecting it, then, if the eyes are placed anywhere between that line and the diameter or that line itself, they will remain invisible. In both cases, no image of either eye will be visible because there will be no intersection between the visual ray and the cathetus of reflection—cases 1 and 5. On the other hand, Euclid goes on to show, if the two eyes are placed between the mirror's surface and the point of bisection on the radius, a virtual image of the eyes will be formed, whereas if the two eyes are placed beyond the diameter, a real image will be formed—cases 2 and 3 (see passage 3.28 below).

²⁶Here Ptolemy is wholly at odds with Euclid, who claims that there can be no image whatever when the cathetus of reflection and the incident ray do not meet somewhere between the mirror and the eye or behind the mirror.

²⁷Passage 3.3 above.

²⁸Notice that these "appearances" are without reference to a center of sight. Theoretically, of course, Z must represent the center of sight, which will then have to lie on the actual object-surfaces.

²⁹This proportionality follows from the fact that angle $HBL = \text{angle } LBT$ within triangle HBT ; see Euclid, *Elements*, VI, 3: "If an angle of a triangle be bisected and the straight line cutting the angle cut the base also, the segments of the base will have the same ratio as the remaining sides of the triangle . . ." For the full theorem, see *Elements*, trans. Heath, vol. 2, pp. 195-197.

³⁰As Ptolemy demonstrates in the next theorem, line TL does not define the actual image but, rather, its rectilinear cross-section, which indicates only its relative size; the image itself will be convex.

³¹This latter claim is in fact treated not in this theorem but in the one following.

³²In other words, we can assume that the entire line-image, from M , through B , to L will be completely continuous and regular in its curvature, as indeed will the line-image from S , through B , to N .

³³The more sharply curved the object, then, the less convex the image until, finally, it forms a straight line or a concave one. Thus, in figure 3.39a, the object is represented by the dotted curve that joins the mirror's surface at points A' and G' and passes through points T and H . The resulting object-curve $A'HTG'$ is sharper than the previous object-curve $AHTG$. As in the previous case, so in this one, since points A' and G' lie on the surface of the mirror, their images will also. So too, the images of T and H will be formed at points N and M respectively, and, as a consequence, the composite image $A'NMG'$ will form a straight line. On the other hand, if the object-curve touches the surface even farther toward points L and K yet still passes through points

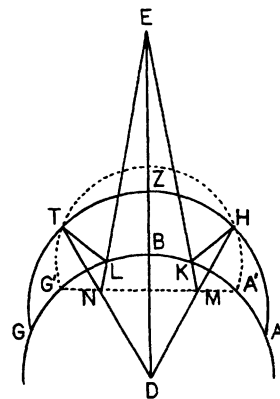


figure 3.39a

T and H, it will be more sharply curved than object-curve A'HTG'. Accordingly, its image will be concave.

³⁴If angle CAZ > angle BZA, then, *a fortiori*, angle PAZ, the complement of angle CAZ, is larger than angle BZA. Therefore, the two lines CA and BZ will intersect in the direction of points A and B.

³⁵Euclid also addresses the problem of image-inversion in isolation from the problems of distortion in propositions 11 and 12 of the *Catoptrics*, pp. 302-306.

³⁶In other words, as in the case of convex mirrors (see passage 3.27 above), so in this one, the shape of the image (i.e., convex, straight, or concave) will depend on how sharply curved arc EBZ of the object is. Although Ptolemy does not specify the fact, it is clear that the center of sight must in this case coincide with the center of curvature D. This, of course, poses an extreme limitation on the resulting conclusions.

³⁷This follows because the resulting ray-couple MLE would not encompass centerpoint D (see note 14 above).

SECTION IV

DIOPTRICS: ANALYSIS OF VISION BY DEFLECTED RAYS

As with reflection, so with refraction, the most fundamental problem has to do with image-displacement—i.e., the object appears to be where in fact it is not. Thus, the analytic goal is to reconcile such apparent position with the object's real position. The analogy between reflection and refraction extends to the supposed corrective: if image-location in reflection is ultimately contingent on a specific relationship (equality) between i (the angle of incidence) and r (the angle of reflection), then presumably the same must hold for refraction, except that the specific relationship will not be one of equality. Accordingly, in the first part of this section we will examine the dynamic model of refraction developed by Hero of Alexandria and Ptolemy and will then look at how, on the basis of both that model and an experimental program, Ptolemy attempted to specify the relationship between i (angle of incidence) and r (angle of refraction) that governs refraction. We will proceed in the second part of this section to Ptolemy's account of how atmospheric refraction affects celestial observations. The third and fourth parts, finally, will be devoted to the analysis of image-location and image-distortion in refraction, the focus here being upon how the apparent place and size of objects viewed through refracting interfaces can be affected not only by the difference in density between optical media but also by the shape of the refracting surface.

1. Observation and Explanation of the Phenomenon

That an oar partially submerged in water appears to break at the water's surface was not only a commonplace observation but also a cause for some wonder in antiquity. As far as optical sources are concerned, the earliest mention of refraction is to be found in the Euclidean *Catoptrics*. The sixth definition in that work asserts that "if some object is placed in a vessel at such a distance [from the viewpoint] that it can no longer be seen, and if water is poured in, then the object placed in the vessel will be seen." Hero of Alexandria also adverts to refraction tangentially in his dynamic account of refraction in the third chapter of the *Catoptrics*, where he treats refraction in terms of imperfect reflection. Having likened reflection to rebound from a hard,

opposing surface, he continues thus:

[4.1] In the case of water and glass, however, not all rays are reflected, for both substances contain interstices and are composed of subtle parts as well as of solid portions. For we see not only our own images but also objects lying beyond in glass and water. Indeed, in marshy waters we see what is at the bottom and through glass we see objects that lie beyond, for some of the rays strike solid portions and, being opposed by them, are reflected, whereas others penetrate through the interstices [and] those rays see what lies beyond. Accordingly, the images that are formed in such bodies are not seen perfectly but, as has been said, some of the rays are deflected through the interstices.

Ptolemy falls back upon much the same dynamic model to account for refraction, which he too associates closely with reflection. This point is clear in the following passage from *Optics*, V, 1-2 (p. 229):

[4.2] There are two ways in which the visual ray is broken. One involves rebound and is caused by reflection from bodies that block the [visual ray's] passage and that are included under the heading of "mirrors." The other way, however, involves penetration and is caused by a deflection in media that do not [completely] block the [visual ray's] passage, and those media are included under the single heading "transparent."

It has been claimed earlier that this sort of bending of the visual ray does not occur [the same way] in all liquids and rare media; what happens instead is that in each one of these [media] the amount of deflection is determined solely by the way in which the medium allows penetration. It has also been claimed that the visual ray radiates rectilinearly, and such rays break only because of an impedance posed by the surfaces separating media of different consistency. It has also been claimed that refraction occurs not only in the passage from rarer and more tenuous to denser media—as happens in the case of reflections—but also in the passage from a denser to a rarer medium.¹

Refraction is not just similar to (or a mere byproduct of) reflection for Ptolemy, however. It is in fact a special case of reflection, as is clear from Ptolemy's brief discussion of the principles that both share in *Optics*, V, 3 (p. 230):

[4.3] At this point we ought to investigate the quantitative relationship between the angles [of incidence and refraction] according to specific intervals. But we should start by discussing the phenomena that such refractions have in common with reflections. First, in either case, whatever is seen appears along the continuation of the incident ray—i.e., along the continuation of the ray that emanates from the eye to the surface at which it is broken—[second, the object appears] on the straight line dropped perpendicularly from the visible object to the surface where the breaking occurs. It

therefore follows that, just as was the case for mirrors, so in this case, the plane containing the broken ray-couple must be perpendicular to the surface where the breaking occurs.

To illustrate the similarity between refraction and reflection in terms of the principles governing image-location, Ptolemy offers an empirical demonstration that has clear echoes of the example offered in the sixth postulate of the Euclidean *Catoptrics* quoted above. Ptolemy's account in *Optics*, V, 5-6 (pp. 230-231) is given below:

[4.4] That this is clear and indubitable we can understand on its own terms by means of a coin that is placed in a vessel called a *baptistir*.² For, if the eye remains fixed so that the visual ray passing over the lip of the vessel passes above the coin, and if water is then poured slowly into the vessel until the ray that passes over the edge of the vessel is refracted toward the interior to fall on the coin, then objects that were invisible before are seen along a straight line extended from the eye to a point higher than the true point [at which the coin lies]. And it will be supposed not that the ray is refracted toward those lower objects but, rather, that the objects themselves are floating and are raised up to [meet] the ray. For this reason, such objects will be seen along the continuation of the [incident] visual ray, as well as along the normal dropped [from the visible object] to the water's surface—all according to the principles we have previously established.

[EXAMPLE V.1] Now, let us suppose that point A [in figure 4.1] is the eye, ZHE the common section of the plane containing the refracted ray-couple and the surface [of the water] in the vessel, and ABD the ray passing over the vessel's lip at B. Let us also suppose that there is a coin at G, which lies toward the bottom of the vessel. Then, as long as the vessel remains empty, the coin will not be seen, because the body of the apparatus at B blocks the visual ray that could proceed directly to the coin. Yet, when just enough water is poured into the vessel so that its surface reaches line ZHE, ray ABH is deflected along line GH, compared to which AH[D] is higher. In that case, then, the coin will appear to be located along the cathetus dropped from point G to EH—i.e., cathetus LKG, which intersects line AHD at point K. Moreover, its image-location will lie on the radial line passing from the eye and continuing rectilinearly to point K, that radial line being higher than the actual ray [HG] and nearer the water's surface; so the image will appear at point K.

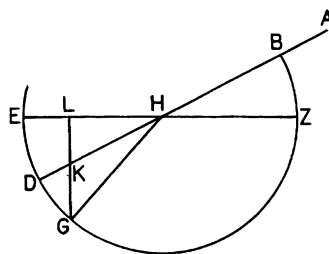


figure 4.1

While reflection and refraction share most of the same principles of image-

location, there is one principle that definitively separates them: whereas in reflection the angles of incidence and reflection are equal, in refraction this is clearly not the case. Yet, as Ptolemy claims at the end of *Optics*, V, 2, even though “this breaking does not take place at equal angles, . . . the angles [of incidence and refraction] do bear a certain consistent quantitative relation to one another with respect to the normals.” What might that “consistent quantitative relation” be? To answer this question, Ptolemy proposes a series of experiments to determine precisely how the angle of refraction varies with respect to the angle of incidence. The first experiment, which is described in *Optics*, V, 7-11 (pp. 231-233), involves refraction from air into water:

[4.5] The amount that the ray is refracted in water below the [original line of] sight is determined according to the following experiment, which is conducted by means of the bronze plaque that we constructed for analysing the phenomena of mirrors.³

[EXPERIMENT V.1] Let circle **ABGD** [in figure 4.2] be described on that plaque about centerpoint **E**, and let the two diameters **AEG** and **BED** intersect one another at right angles. Let each of the [resulting] quadrants be divided into 90 equal increments. At the centerpoint let a small marker of some color or other be attached, and let the plaque be stood upright in the small vessel [discussed in the previous experiment]. Then let a suitable amount of water that is clear enough to be seen through be poured into that vessel, and let the marked-off plaque be placed erect at right angles to the surface of the water. Let all of semicircle **BGD** of the plaque, but nothing beyond that, lie under water, so that diameter **AEG** is normal to the water’s surface. From point **A**, let a given arc **AZ** be marked off on either of the two quadrants that lie above the water. Furthermore, let a small, colored marker be placed at **Z**.

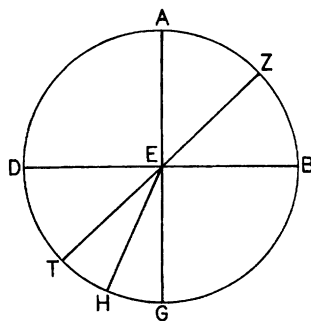


figure 4.2

Now, if we line up both markers at **Z** and **E** along a line of sight from either eye so that they appear to coincide, and if we then move a small, thin peg along the opposite arc **GD** under water until the end of the peg, which lies upon that opposite arc, appears to lie directly in line with the two previous markers, and if we mark off the portion of the arc **GH** that lies between **G** and the point at which the object would appear unrefracted, the resulting arc will always turn out to be smaller than **AZ**. Moreover, if we join lines **ZE** and **EH**, angle **AEZ** > angle **GEH**, which cannot be the case unless there is refraction—that is, unless ray **ZE** is refracted toward **H** according to the

excess of one of the opposite angles over the other.

Furthermore, if we place our line of sight along normal AE , we will find the image directly opposite along its rectilinear continuation, which will extend to G ; and this [radial line] undergoes no refraction.

In the case of all the remaining positions, when arc AZ is increased, arc GH in turn will be increased, and the refraction will be greater. When arc AZ is 10 degrees out of the 90 into which quadrant $[AB]$ is divided, then arc GH will be around 8 degrees.⁴ When AZ is 20 degrees, then GH will be 15.5. When AZ is 30 [degrees], then GH will be 22.5. When AZ is 40 [degrees], then GH will be 29. When AZ is 50 [degrees], then GH will be 35. When AZ is 60 [degrees], then GH will be 40.5. When AZ is 70 [degrees], then GH will be 45.5. And when AZ is 80 [degrees], then GH will be 50.

The second experiment, which involves refraction from air to glass, is described in *Optics*, V, 14-18 (pp. 234-236):

[4.6] [EXPERIMENT V.2] Let a clear glass semicylinder, represented by arc TKL [in figure 4.3] be made in accordance with the semicircular section of the round plaque, but let its diameter be smaller than the diameter of the aforementioned bronze plaque. Then let its base be attached to the plaque so that both are completely joined. Let its center be E , and let its diameter TL coincide with [the plaque's] diameter BD , and let AE be perpendicular to the flat surface of the glass. Therefore, every line drawn from point E to arc BGD and to arc TKL will be normal [to those arcs].

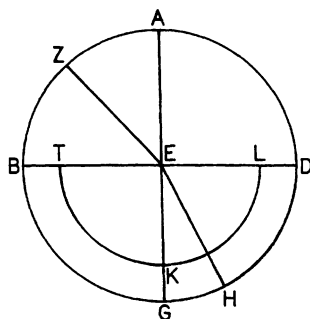


figure 4.3

Accordingly, if we set up this experiment as we did before, and if we make a small mark on the mid-point of the semicylinder's surface where it[s axis] touches point E , and if we look with either eye along line AE toward the edge of the glass and move a marker on the arc $[BGD]$ opposite this arc $[BAD]$ until it appears in front of it, it will be found to lie on G itself. For line AEG is normal to both TEL and TKL . And if we move our eye until it lies directly in line with this position, and if we look along line GE so that the marker that has been moved along the arc lies directly in line with GE , then that marker will be situated on line EA . For the same reason, moreover, there will be no refraction in the passing of the ray [orthogonally] from glass into air.

But if we take some given arc AZ from point A and draw line ZE , coloring it black, and then if we sight along this line until the marker, which is moved behind the glass, appears to fall in line with it, and if we mark the place—e.g., point H —where we found it so that the black color coincides

with EH, then we will also find in this case that angle $AEZ >$ angle GEH . We will also find that the angular difference is greater than the angular difference in water, where the arc [measuring incidence] was the same.

If, moreover, we station our eye at point H, which is opposite point E, and sight from point H along HE, [both of the points] E and Z will appear to coincide on one and the same line of sight. And since there appears to be a refraction of the ray in this situation, it is necessary that, whether the ray passes from air into glass, as represented by ZE, and is refracted along EH, or whether it passes from glass to air, as represented by HE, and is refracted along ZE, the refraction takes place toward T.⁵ And since the normals dropped from E to TKL are the same, rays [that pass along them] are not refracted, whether they pass from E to K or from K to E.

In addition, if we now analyze the amount of refraction for each of [the previous angular] positions, we will find that, when the eye is placed at the same angular distances as before and when the angle measured from point E (i.e., the angle [of incidence] formed by normal AE and ray EZ) is 10 degrees of the ninety ascribed to the circle's quadrant, then the [resulting] angle [of refraction] GEH will measure nearly 7 degrees. When the first angle is 20 degrees, the second will be 13.5. When the first is 30 [degrees], the second will be 19.5. When the first is 40 [degrees], the second will be 25. When the first is 50 [degrees], the second will be 30. When the first is 60 [degrees], the second will be 34.5. When the first is 70 [degrees], the second will be 38.5. And when the first is 80 [degrees], the second will be 42.

The third and final experiment involves refraction from water to glass. It is described in *Optics*, V, 20-21 (pp. 236-237):

[4.7] [EXPERIMENT V.3] Let the glass semicylinder [in figure 4.4] be attached to the bronze plaque and let it be set up so that its center coincides with the center of the plaque. Let point E be colored somewhat, and let the plaque be set up in the basin at right angles to the water's surface, which coincides with the plaque's [horizontal] diameter, and let the curved surface TKL of the glass [semicylinder] be arranged such that it lies above [the water's surface]. Then, let water be poured into the basin just until line TEL of the semicylinder coincides with the water's surface. Within the rarer medium (i.e., water) let some arc GH be marked off, and again let it be 10 degrees. Let some thin colored marker be placed at H, and sight with either eye [along HE] until marker Z, which is moved on arc AB, appears to lie in line with point H and the colored spot at E. With things so disposed, let the two lines EH and EZ

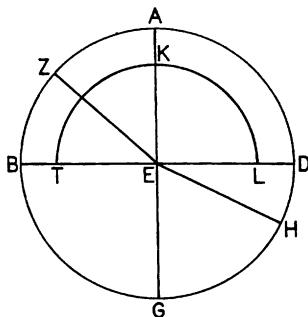


figure 4.4

be drawn.

If, then, we take the angle formed in the denser medium (i.e., in glass), that angle being designated on arc AB , we will find by means of this experiment that, when the angle in water as measured from the normal—namely, angle GEH —is ten degrees of the ninety ascribed to a right angle, the [resulting] angle in the glass—namely, angle AEZ —is nearly 9.5. When the angle in water is 20 [degrees], the angle in glass will be 18.5. When the former is 30 [degrees], the latter will be 27. When the former is 40 [degrees], the latter will be 35. When the former is 50 [degrees], the latter will be 42.5. When the latter is 60 [degrees], the former will be 49.5. When the latter is 70 [degrees], the former will be 56. And when the former is 80 [degrees], the latter will be 62.⁶

With the experimental evidence fully presented, Ptolemy is prepared to make some generalization about refraction in *Optics*, V, 31-35 (pp. 242-244):

[4.8] It is, however, possible to formulate in such a way a general claim about refraction on the basis of previously established points. We put it thus: the amount of refraction is the same whichever the direction of passage; the difference is one of kind [rather than of degree]. For in passing from the rarer to the denser medium, the ray inclines toward the normal, whereas in passing from the denser to the rarer medium, it inclines away from the normal.⁷

[EXPERIMENT V.4] In fact, if we set up the plaque as before and assume that diameter BD [in figure 4.5] lies on the interface between the two different media, and if we draw normal AEG as well as the refracted ray-couple ZEH inclined toward the normal, with which it forms angle GEH , then the path of refraction remains one and the same. In fact, when the visual ray passes through point E , and the eye is stationed at point Z , the line [of sight] after refraction—i.e., line EK —inclines toward the normal according to its continued passage [along EH] while the visible object is seen along the rectilinear continuation $[EK]$. But if the eye lies at point H , and EZ lies within the rarer medium delimited by ABD , then, after refraction, line EL will take an opposite tack from that previously specified, inclining away from normal AE [along EZ] in such a way that it lies farther out [from the normal] than would the visual ray if it were to continue in a straight line.

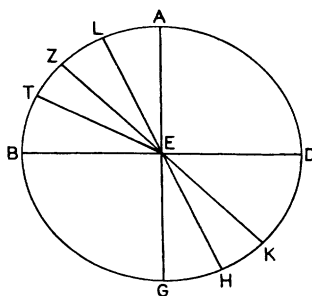


figure 4.5

Furthermore, when the media and the angles differ from one another by a significant amount, the difference [between the angles of incidence and refraction] increases as the density of either of the media grows. Indeed, if

we assume that arc **BAD** lies in the rarer medium and arc **BGD** in the denser, and if we take angle **AEZ** as it is represented, then, when the medium within section **BGD** becomes denser than it previously was, the difference between [angle **AEZ** and] angle **GEH** will vary with the difference in density between the two media. In fact, when angle **AEZ** is 30 degrees in air, angle **GEH** in water will be nearly 22.5 degrees, whereas in glass it will be nearly 19.5 degrees. And in this latter case the refraction and difference in angles measured from the point at the top [of the normal] will be greater, because the substance of glass is denser than the substance of water.

So too, if we suppose that the refraction of another of the visual rays takes place at some other arcal distance [than **AZ**] from normal **AE**—e.g., along ray-couple **TEK**—then $AT : AZ > GK : HG$. By alternation, $AT : GK > AZ : GH$. By separation, $TZ : AZ > KH : GH$. And [by alternation,] $TZ : KH > AZ : GH$.⁸ Furthermore, we can determine particular cases on the basis of the refractions as we measured them if we take the resultant numbers and, on their basis, investigate particular measurements of this sort, substituting the numbers so derived for the two arcs **AZ** and **AT**.

But someone might object by demanding an explanation for why, in light of the first principles laid out concerning normals and the appearance of the visible object along the straight continuation of the [incident] visual ray, the sort of breaking that we have discussed is similar to the breaking that occurs in mirrors, whereas the angular relationships are not, for in the case of refraction the equality of angles is not conserved. It will be seen in response that this must be the case according to what we have explained, on which basis an even more marvelous fact will be apparent: namely, the course of nature in conserving the exercise of power.⁹

Suffice it to say that the governing principle of refraction formulated by Ptolemy in the penultimate paragraph is extraordinarily weak, amounting to little more than the generalization that, in refraction from a given rarer to a given denser medium, as i increases, so does the difference between successive r 's—which is to say that, as i decreases, r and i approach equality. There is, in fact, a specific mathematical law implicit in Ptolemy's tabulations, but its proper formulation in algebraic terms would have been beyond Ptolemy given the limitations of mathematical notation in his day.¹⁰

2. *Practical Application: The Problem of Atmospheric Refraction*

The Effect on Observed Position: According to the cosmology of Ptolemy's day, the universe can be thought of in terms of concentric spheres, at the center of which lies the earth with its oceanic covering. That sphere is surrounded by a spherical sheath of air, which is enveloped in its turn by fire. Beyond that lies ether, the stuff of the celestial realm, whose inner edge is demarcated by the lunar orbit. Ptolemy's explanation of atmospheric refrac-

tion and its effect on celestial observation is to be found in *Optics*, V, 23-30 (pp. 238-242):

[4.9] Furthermore, it is possible for us to realize that, at the surface between air and ether, there is a refraction of the visual ray according to the difference in density between these two media. The resulting phenomena are as follows:

We notice that [celestial] bodies that rise and set tend to incline toward the north when they are near the horizon and are measured by an instrument for measuring the stars.¹¹ For, when they lie to the east or west [i.e., at rising or setting], the circles drawn through them parallel to the equator are nearer to the north than the circles drawn through them when they are in the middle of the sky, and the more they approach the horizon, the more they are inclined to the north. Moreover, the distance from the north pole of the stars that are always visible [i.e., that do not set] will always be less when they lie on the meridian toward the horizon. For when they are on the meridian in a location that is closer to the zenith, the circle at that location that is parallel to the equator becomes larger, whereas in the previous position it becomes smaller.¹² This is a result of the refraction of the visual flux at the surface that separates the air and the ether, a surface that must be spherical, its center being the common center of all the elements, which is the center of the earth.

[THEOREM V.1] First, let E [in figure 4.7] represent the viewer's zenith, and let one of the great circles on the spheres of the aforesaid elements cut the earth along AB. Let GD be a great circle on the interface between the air and ether, and let EZ be a great circle passing through a given star, and let the center of all the circles be point H. Let line EAH be drawn. Let point A be the eye and line ADZ a line coinciding with the common section of the [plane of the] horizon and circle GD. In addition, let DT be normal

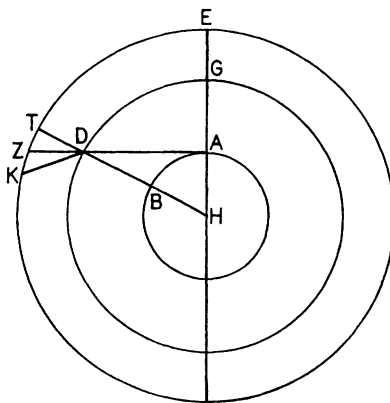


figure 4.7

to the circle[s]. Let us suppose that **ADK** is a visual ray refracted along **KD** at point **D**, and let the star lie at point **K**.

Since the visual ray is refracted at the interface toward a position away from point **E** according to the normal dropped at equal angles to the refracting surface, angle **KDT** [of refraction], which lies in the subtler medium, will be greater [than angle **ADB** of incidence]. Thus, the star will be seen from point **A** along line **ADZ**, and its [apparent] distance from the zenith will be less than its true distance, for it will be seen along arcal distance **EZ** instead of arcal distance **KE**. Therefore, the higher its distance [toward the zenith], the smaller the difference between the star's apparent and true location. And if the star is at **E**, there will be no refraction, because there is no breaking of the visual ray that passes from point **A** to point **E**, for in that case it will be normal to the surface of refraction.

[THEOREM V.2] With these points established, let **ABG** [in figure 4.8] represent the circle of the horizon and **AEZG** the semicircular arc of the meridian that lies above the earth. Let point **E** represent the zenith and point **Z** the apparent pole of the heavenly sphere [of fixed stars]. Let **BHD** be the arc of a line above the earth's surface that is parallel to the equator and that passes through certain stars.¹³ Let **T** be a star that lies on this line near the horizon, and let **KETL** be the semicircular arc of the circle that passes above the earth through the zenith and through star **T**.

Accordingly, since the star appears to lie nearer the zenith than it truly is

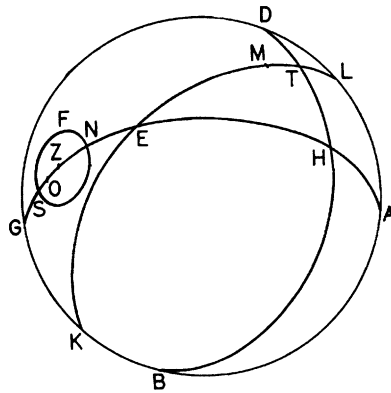


figure 4.8

when it is near the horizon, and since the divergence in its apparent from its true position is measured on the great circles passing through the points on the horizon, the point where the star that is [actually] at **T** appears will lie between **E** and **T**, such that it appears at point **M**. Moreover, the line parallel to the equator and passing through point **M**, will lie higher to the north than the line parallel to the equator and passing through point **T**, which at

our particular latitude is inclined toward the north. And when the star rises to position H, it reaches a point where the visual ray is refracted without any perceptible difference between apparent and true location.¹⁴

Likewise, if we suppose Z to represent the north pole and if we draw one of the circles parallel to the equator that is always visible [at our latitude], e.g., circle NSF, then, when the star lies at point S on this circle, it will appear closer to point E, which lies at the zenith, and it will seem to lie at point O. But when the star is at point N, then there is no difference, or only an imperceptible one, between apparent and true location.¹⁵ And therefore, when a star approaches the horizon in its revolution, its distance from the north pole of the [celestial] sphere appears to be smaller [than it really is]; but when it approaches the zenith in the course of its revolution, that apparent distance seems larger, for arc ZN will be larger than arc OZ.

It has thus been demonstrated how stellar observation must be affected by the refraction of the visual ray. It would also be possible for us not only to examine the degree of such refractions, but also to analyze such refraction in the case of certain [celestial] bodies whose distance is given—e.g., the sun and the moon—and to determine the degrees [of refraction] toward the horizon as well as the amount by which the refraction of the visual ray shifts the apparent position upward if the distance of the interface between the two media [i.e., air and ether] were known.¹⁶ But, although this distance lies nearer than the earth to the lunar sphere, where the ether stops, it is not known whether the [refractive] interface lies at the same distance as the aforesaid surface, or whether it lies nearer the earth, or whether it lies farther from [the aforesaid] surface.¹⁷ Therefore, it is impossible to provide a method for determining the size of the angles of deviation that occur in this sort of refraction.

Despite Ptolemy's clear articulation of the problem of atmospheric refraction, it was not until the very end of the sixteenth century that astronomers made serious efforts to determine the effect of that refraction on celestial observations.

3. *Image-Location as a Function of the Shape of the Refracting Surface*

As in reflection, so in refraction, image-location is affected by the shape of the refracting interface. Ptolemy offers an analysis of this effect for plane, spherical convex, and spherical concave refracting surfaces in *Optics*, V, 56-63 (pp. 251-253):

[4.10] At this point, we want to demonstrate how the cathetus dropped from the visible object to the aforesaid [refracting] surface intersects the refracted ray.

[THEOREM V.8] To start with, let the [refracting] surface be plane,

and let straight line **ABG** [in figure 4.10] represent the common section of that surface with the plane containing the refracted ray-couple. Let **D** be the eye and **DBE** an oblique ray. Draw normal **ZBH** to point **B** on **AB**, and let **DB** be refracted toward the normal along **BT**, on the one hand, and away from the normal along **KB**, on the other. Let the visible objects lie at **T** and **K**, and from them let the two catheti **KL** and **TG** be dropped to **BG**. They invariably intersect line **DB**, for angle **ABE** is larger than a right angle, while the angles at **G** and **L** are right.

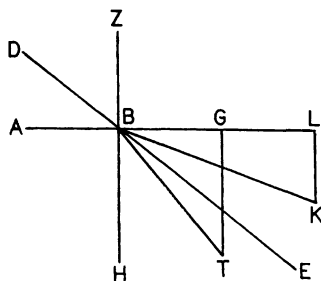


figure 4.10

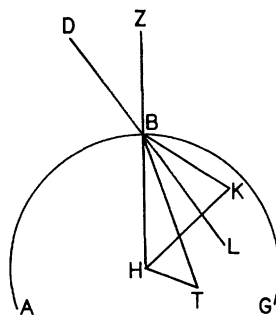


figure 4.11

The unstated but obvious conclusion of this demonstration is that the image of **T** will appear at the juncture of **BE** and **GT**, whereas the image of **K** will appear where the extensions of **LK** and **BE** intersect.

[THEOREM V.9] Now, let the surface of refraction [in figure 4.11] be spherical, [let its center be **H**,] and let its convex surface face the eye. Let **D** be the eye, and let [incident ray] **DBL** be drawn.

Accordingly, catheti, such as **KH**, that are drawn to any of the visible objects placed on **KB** between **Z** and **H**¹⁸ invariably intersect **BL**. However, catheti, such as **TH**, that are drawn to such objects and that intersect **BT** and **ZH**, will sometimes be parallel to **DBL**, sometimes intersect it in the direction of **T** and **L**, and sometimes diverge from it. For it is possible for angle **HTB** to be at times equal to angle **LBT**, at times greater than it, and at times smaller than it.

The point of this demonstration is that the image of an object immersed in a rarer medium will always be seen, because the cathetus dropped from the object to the center of curvature will always intersect the continuation of the incident ray. On the other hand, if the medium within **ABG** is denser so that the visual ray refracts toward normal **BH**, an image may or may not appear, depending upon whether the cathetus **TH** intersects the continuation of incident ray **DB** or not.

[THEOREM V.10] Now, let the concave surface face the eye, which lies

at any one of the two positions [between the center of curvature and the surface or beyond the center of curvature]. Let ray DBL [in figures 4.12a and 4.12b] be drawn, and let the visible objects be assumed to lie on BT and KB.

According to both positions for the eye, then, the cathetus TG dropped

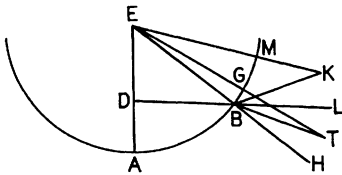


figure 4.12a

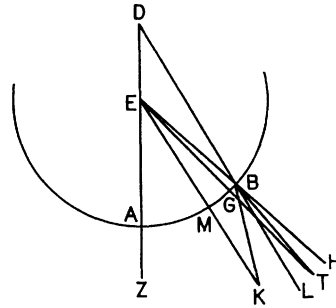


figure 4.12b

from the visible object on BT will intersect ray BL. But the cathetus KM dropped from any of the visible objects on KB will sometimes be parallel to line DBL, sometimes intersect it on the side of K and L, and sometimes diverge from it. For angle KBL can sometimes be equal to angle BKE, sometimes larger than it, and sometimes smaller than it, depending on the distance [along BK from the point of refraction] of the object seen at K.

In this case, moreover, it is possible that the aforementioned lines—i.e., the refracted ray and the cathetus dropped from the visible object to the refracting surface—do not meet and [therefore] that there will be no determinate image-location. This is possible when the refracting surface is spherical and its convex surface faces the eye, which lies in the rarer medium, or when the concave surface faces the eye, which lies in the denser medium.

In every other case, though, the lines we have designated always intersect. But when they do not intersect, the visual faculty is affected in the same way that it is in the case of mirrors in regard to the intersection except that the location where the image is formed will not be determinate but is shifted to the common intersection of the normal and the refracting surface, and it takes on the image of the visible object and will coincide with it in location and in the transparent medium.

From the above account it is clear that Ptolemy is constrained by his visual model to assume that, whenever the ray actually reaches an object, the object must somehow be seen. As a result, he is forced to temporize by supposing that, at times, the “image” is confused with the surface of refraction and thus takes on its visible qualities.

4. Size-Distortion

The most obvious form of image-distortion in refraction involves the magnification of objects when they are viewed through water or glass. This effect was recognized at least a century before Ptolemy wrote his *Optics* and was drawn on to explain why celestial bodies appear so much larger at the horizon than they do at zenith. The explanation itself rests on the assumption that the vapor-laden atmosphere acts like a vessel of water through which celestial bodies appear larger than they are. This magnification-effect is most pronounced at the horizon, where the angle of incidence for the visual ray is greatest and, consequently, where the depth of atmosphere traversed by the ray is greatest. Thus, as Ptolemy summarizes the explanation in *Almagest*, I, 3¹⁹

[4.11] The apparent increase in [the sizes of celestial objects] at the horizon is caused, not by a decrease in their distances, but by the exhalations of moisture surrounding the earth being interposed between the place from which we observe and the heavenly bodies, just as objects placed in water appear bigger than they are, and the lower they sink, the bigger they appear.

Later, in his *Planetary Hypotheses*, and even later in the *Optics* (III, 59 [p. 151]), Ptolemy shows that he has become fully aware that the apparent enlargement of celestial objects at the horizon (the so-called Moon Illusion) is *only* apparent and thus due to psychological rather than physical causes.

Although Ptolemy's attempt in the *Almagest* to explain the Moon Illusion on the basis of refractive distortion was badly misguided, that explanation nonetheless reveals Ptolemy's awareness that refraction causes size-distortion. By the time he composed the *Optics*, in fact, he had a clear understanding that such size-distortion is due to image-displacement and, moreover, that such displacement is a function of the difference in density between optical media. Accordingly, he starts in *Optics*, V, 70-72 (pp. 255-256) with the case of a point-object viewed in a denser medium, such as water:

[4.12] [THEOREM V.11] Let straight line ABG [in figure 4.13] be the common section of the water's surface and the plane containing the refracted ray-couple. Let D be the eye, let ray DB be drawn, and let normal ZBH pass through point B. Let ray DB be refracted along BT toward the normal, as happens when the eye lies in the rarer medium. Then, let [cathetus] TKG be dropped normal to AG from point T.

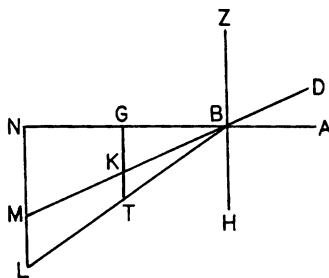


figure 4.13

Accordingly, an object that lies at point **T** will be seen at **K**. And since lines **DB** and **BT** taken together are longer than line **DBK** (for angle **BKT** is obtuse), the distance of the visible object's image will be less than that of the object itself.

If, moreover, we draw **BTL** and place the visible object at **L**, and if from that point we draw [cathetus] **LMN** normal to **AG**, the image of **L** will lie at **M**, and the image that appears at point **M** will lie farther away than the image that appears at point **K**. But it will not just lie farther away; it will do so according to a particular mathematical relationship that is proper to it. Thus, since **TG** and **LN** are parallel and lie on a plane, $MN : KG = MB : KB = LB : TB$. And so, $MN : KG = LB : BT$.

Ptolemy then turns in *Optics*, V, 74 (p. 256), to the opposite case: a point-object viewed in a rarer medium, such as air.

[4.13] [THEOREM V.12] For if we assume refracted ray **BTL** [in figure 4.14], which forms an angle [of refraction] **LBH** greater than angle [of incidence] **DBZ**, then the ratios of the distances will remain constant to one another. For the ratio of distance **L[B]** to distance **T[B]** will be the same as the ratio of distance **M[B]** to distance **K[B]**. But in each case, the relationship is the inverse [of that in the previous demonstration]. For the location of **T**, which represents the visible object itself, will be nearer than [that of] its image **K**, just as the location of **L** will be nearer than [that of] its image **M**.

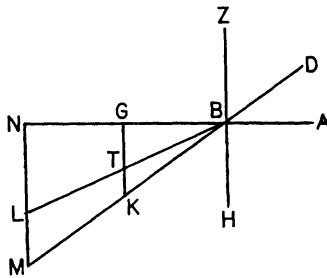


figure 4.14

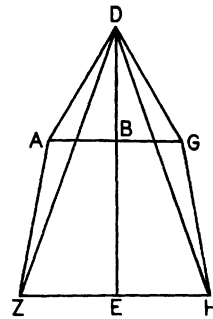


figure 4.15

Moving from point-objects to lengths in *Optics*, V, 76-78 (pp. 257-258), Ptolemy explains the overall magnification caused by refraction in terms of the size of the visual angle under which the image is seen:

[4.14] [THEOREM V.13] Let the two rays **DA** and **DG** [in figure 4.15] be drawn from the eye at **D** to line **GA**, and let them flank the normal **DBE**. Then let them be refracted so as to flex apart from the normal along **AZ** and

HG, and let them comprehend some magnitude represented by the line connecting the two endpoints of magnitude ZEH. Finally, let lines DZ and DH be joined. It is therefore evident that angle ADG > angle ZDH, because the refracted rays lie farther from the normal [than the continuation of the incident rays along DZ and DH]. And [so], ZH will be seen under a larger angle [than otherwise], assuming that the distance and disposition [of the magnitude] remain constant.

For this reason, then, objects that are submerged in water must invariably appear larger than they would if they were observed according to direct vision at the same distance and in the same disposition.

[THEOREM V.14] But, assuming that the situation is the reverse of that just specified, let us suppose that rays DA and DG [in figure 4.16] are refracted away from the normal, as happens when the eye is situated in the denser medium. Thus, if we join lines DZ and DH, angle ZDH > angle ADG. And so the actual object must appear larger than its image.

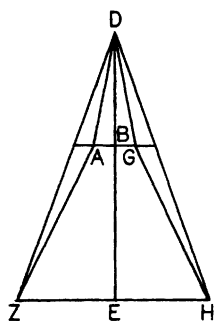


figure 4.16

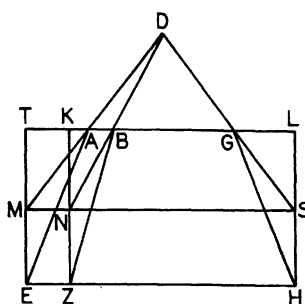


figure 4.17

Having determined that the amount of magnification is due to the increased size of the visual angle under which the image is seen, and having also determined that the fundamental cause of such magnification is the upward displacement of the image, Ptolemy concludes in *Optics*, V, 80-82 (pp. 258-259), that, in the case of refraction through a plane interface, the image appears to be precisely the same shape as its generating object:

[4.15] [THEOREM V.15] . . . If ray-couple DAE [in figure 4.17], ray-couple DBZ, and ray-couple DGH are each refracted toward magnitude EZH, and if catheti of refraction ET, ZK, and LH are dropped, while lines DAM, DBN, and DGS are connected, then the image of line EZH will lie on the line passing through points M, N, and S. Also, $TM : KN = ET : ZK$, while $KN : SL = KZ : LH$, and $TM : SL = ET : LH$.

The same will hold as well for every type of disposition. For if segments TE, KZ, and LH define the observed line EZH as a straight line, then seg-

ments **MT**, **KN**, and **SL** define the observed [image]-line as a straight line. On the other hand, if the former segments define the observed line as convex, the latter segments define the observed image as convex. Finally, if the former segments define a concave line, the latter define the observed image as concave.

As is represented by the figure [4.18], moreover, this is the case whether we suppose the refraction to occur in the direction of, or away from, the normal.

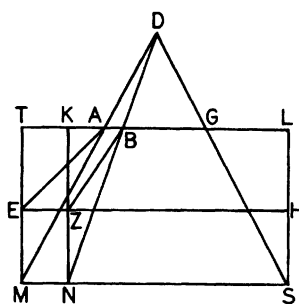


figure 4.18

Even the most cursory glance into a moderately large flat-bottomed container filled with water will show the basic conclusion of this theorem to be false: since the bottom in fact appears concave, it is clear that the image does not have the same shape as its generating object. Indeed, Ptolemy's procedure for determining the image-location of point-objects should have led him to conclude that M is projected somewhat higher than N, so that, all told, the image of that portion of EH left of midpoint will incline constantly upward as the line of sight moves ever outward. The source of Ptolemy's error, of course, lies in the false precision of the three proportionalities, $TM : KN = ET : ZK$, $KN : SL = KZ : LH$, and $TM : SL = ET : LH$. The simplicity of the resulting model apparently seduced Ptolemy into jumping to a conclusion that was inconsistent with his own principles.

NOTES

¹This paragraph is rife with implications. First, it is evident that Ptolemy sees a fundamental relationship (direct proportionality?) between the density of a given transparent medium and the amount of impedance it poses to the passage of visual rays. Second, Ptolemy seems to believe that the amount of deflection or refraction of such rays is fundamentally related (in direct proportionality?) to the amount of impedance posed by the given medium. Thus, we are led to conclude a fundamental relationship (direct proportionality?) between density and amount of refraction. The problem, of course, is how to gauge “density” in this account. Should we somehow take it as physical, so that it measures the relative amount of matter contained in a given volume? If so, what is the relative density, for instance, of glass to water? Worse, why are so many opaque objects less dense than transparent ones? In view of these kinds of issues, we should be chary of assuming that Ptolemy has a true model for refraction in mind. Perhaps all he means to convey is a vague sort of analogy.

²From later discussion in book V, we know that the *baptistir* to which Ptolemy refers consists of a hollow semicylinder whose ends are closed off so that it can hold water.

³See passage 3.3 above.

⁴The qualifier “around” in this context suggests that Ptolemy was self-consciously rounding his results off—to the nearest half degree as it turns out.

⁵Here, for the first time, Ptolemy explicitly articulates the principle of reciprocity: i.e., no matter the direction of passage in refraction (or, for that matter, reflection), the ray-couple remains fixed, with the incident and refracted rays interchangeable.

⁶It is worth pausing at this juncture to discuss the methodological implications of these experiments. To facilitate that discussion I will tabulate the results for all three refractions, i being the angle of incidence and r being the observed angle of refraction, both measured with respect to the normal:

air to water		air to glass		water to glass	
i	r	i	r	i	r
10°	8.0°	10°	7.0°	10°	9.5°
20°	15.5°	20°	13.5°	20°	18.5°
30°	22.5°	30°	19.5°	30°	27.0°
40°	29.0°	40°	25.0°	40°	35.0°
50°	35.0°	50°	30.0°	50°	42.5°
60°	40.5°	60°	34.5°	60°	49.0°
70°	45.5°	70°	38.5°	70°	56.0°
80°	50.0°	80°	42.0°	80°	62.0°

The most obvious characteristic of all three sets of tabulations is that they are not “raw;” clearly, Ptolemy has rounded them off to the nearest half-degree, presumably because of the limits of observational accuracy. Only a little less obvious is the fact

that they all follow a pattern. Let us take the tabulations for refraction from air to water as an example. First, in comparison to the proper values for the angles of refraction as computed on the basis of the sine law (sine i : sine r is constant for refraction between any two given media), Ptolemy's values are relatively accurate for all values of i except for the last value—i.e., $i = 80$. Second, the structure of the tables indicates that, along with recognizing the inexactness of his brute observations, Ptolemy had a specific algorithm in mind: namely, that the progression of values for r depends on constantly decreasing increments, the rate of decrease being constant at half a degree. If, however, r were to vary constantly with i , then the progression of values would be as follows: when $i = 10$, $r = 8$; when $i = 20$, $r = 16$; when $i = 30$, $r = 24$, etc., the difference in successive values being 8. As tabulated by Ptolemy, though, the successive values are continually decreased by half a degree so that the resulting values for r become 8; $8 + 7.5 (= 8 - .5) = 15.5$; $15.5 + 7 (7.5 - .5) = 22.5$, etc. The numbers in italics represent what are called "first differences." Thus, as i increases from 10 to 80, the first differences decrease continually. But in fact those differences decrease by a constant amount of half a degree, which are represented by the numbers in bold-face type. These increments of half a degree are called "second differences." Hence, at bottom, Ptolemy was adjusting his results according to an algorithm of constant second differences, an algorithm that was commonly used in late Babylonian astronomy for generating astronomical tables. In short, Ptolemy was drawing upon his astronomical background to address this particular computational problem.

⁷Here, and in the following paragraph, Ptolemy is simply generalizing the claim about reciprocity made above in passage 4.6, paragraph 4.

⁸In essence, this entire series of ratios expresses the fact that the difference between successive i 's increases more slowly than the difference between successive resulting r 's, or, to put it another way, r approaches i as i approaches 0° .

⁹Unfortunately, this extraordinarily suggestive point is not followed up in the remainder of the text as we now have it. It is tempting to think that Ptolemy has in mind some sort of principle of conservation, either dynamic or kinematic, that governs the angular relationship between i and r .

¹⁰The precise law implicit in Ptolemy's data can be formulated thus: if $n = i/10$, $R = n \times$ the value of r for $i = 10^\circ$, and $d_2 =$ the second difference $= .5^\circ$, then for any given i , $r = R - n(n-2)d_2/2$.

¹¹Presumably what is being referred to here (by Ptolemy or one of his translators, whether Arabic or Latin), is an astrolabe, or something akin to it, that measures the latitude of celestial bodies.

¹²According to Ptolemy, then, the apparent latitude of the stars changes constantly as they rise higher in the sky toward the zenith in their nocturnal round, even though their actual latitude remains constant. For instance, let line EW [in figure 4.6] represent the horizon-line, with E marking the eastward side and W the westward side. Let the dotted line pointing toward N (= North) represent the meridian line, and let arc AB represent the path of a given star. The observer's line-of-sight will therefore be along the perpendicular dropped from a point above the plane of the diagram to the intersection of horizon-line EW and the meridian line. Now, accord-

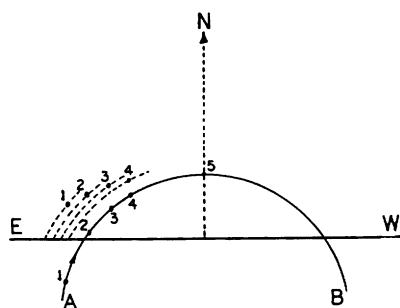


figure 4.6

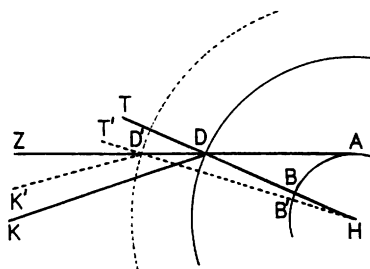


figure 4.9

ing to Ptolemy's account, while the star is still below the horizon at position 1, it will appear above the horizon at counter-position 1 because of atmospheric refraction, which causes the star, being in a rarer medium (i.e., ether), to appear to "float" above its actual position (see passage 4.4 above). Thus, the projection of its apparent latitude or orbit (along the dotted line) describes a circle lying above—i.e., to the north of—its actual circle of latitude. Likewise, as it progresses along its actual orbit from 2 to 5, refraction causes the apparent position to lie farther north than the real one, but the amount of refraction diminishes continually as it rises higher in the sky. Thus, the dotted circles of apparent latitude or orbit constantly approach the actual circle of latitude or orbit until the star reaches zenith (point 5), where no refraction occurs and where the apparent orbit has reached as far south as possible. Then, and only then, will the projected circle of latitude or orbit coincide with the actual one. Moreover, the farther north the circles of latitude lie, the smaller they necessarily are in comparison to the great circle of the equator.

¹³In short, BHD represents a line of latitude.

¹⁴In other words, point M represents one of the intermediate positions (2-4) of the star in figure 4.6, note 12 above, whereas E represents position 5 in that figure.

¹⁵That is, when the star is at S and therefore closest to the horizon in relation to zenith-point E, the ray that reaches it will be maximally refracted. Consequently the apparent location of the star will be shifted northward as much as possible. On the other hand, when the star is at N, it is closest to zenith-point E and, therefore, to the point at which refraction is minimal. Hence, its northward shift will be the least possible.

¹⁶The amount of refractive displacement in this case will depend upon two basic factors: 1) the actual refractive power of the denser medium (air) relative to that of the rarer one (ether), and 2) the depth of the denser medium. Ptolemy seems to be citing the distance of the bodies in question as another factor, but in fact, the actual eye-to-object distance has virtually no effect on refractive displacement (perhaps by "distance" he means actual latitudinal distance with respect to the equator at any given point along the known orbit of the body). All of this can be graphically explained by recourse to figure 4.9, which is adapted from figure 4.7 above. Let arc ABB' represent an arc on the earth's surface, with H as its center and A as the viewpoint. Let the two arcs through D and D' represent two possible interfaces between air and ether, and let HBDT and HB'D'T' be the respective normals to those inter-

faces. Finally, let ADD'Z represent the horizontal line-of-sight. From the diagram, then, it is clear that, for the nearer interface, a given body K will be displaced upward by refraction farther than a body K' viewed through the farther interface. It is also clear that any body along line DK will have the same angular displacement, no matter its actual distance along that line. Note, finally, that the farther out the refractive interface lies, the smaller the amount of refraction; indeed, there is a point at which the refraction is minimized to imperceptibility.

¹⁷This passage is confusing, because it seems to imply the possibility that the air-ether interface might lie beyond the lunar sphere, which is a patent absurdity given his earlier recognition that the moon is one of the bodies of known distance whose apparent location is affected by refraction. According to Aristotelian cosmology, however, the air-ether interface cannot possibly lie below the lunar sphere, since that sphere defines the lowest limit for ether. Also, according to that same cosmology, there is a sphere of fire separating air from ether, yet that sphere goes unmentioned.

¹⁸This specification, "between Z and H," is at best irrelevant and at worst incoherent.

¹⁹From G. J. Toomer, trans., *Ptolemy's Almagest* (New York: Springer Verlag), p. 39.

SECTION V

ANALYSIS OF THE RAINBOW AND OF BURNING MIRRORS

As mentioned in the introduction, the analysis of the rainbow seems not to have been integral to the science of optics in antiquity—at least as that science is currently represented in the rather paltry textual tradition that is currently available. The only significant study of the rainbow that survives from that period, in fact, is Aristotle's account in the *Meteorology*, which is a work devoted not to optics but to what were, or were presumed to be, atmospheric phenomena (including comets). Whether the survival of Aristotle's rainbow-analysis is a function of its overwhelming authority—by virtue of which it buried all competition—or whether that survival is a mere historical accident is impossible to determine with certainty. We do have suggestions of other accounts in antiquity, including one by Ptolemy.¹ Given the fragmentary and questionable nature of these suggestions, though, our discussion of the rainbow in the first part of this section will necessarily focus on Aristotle.

Whereas it is not clear why the study of rainbows was considered peripheral by ancient optical theorists (after all, the rainbow represents a visual effect *par excellence*), the study of burning mirrors is another matter entirely. For burning mirrors have nothing to do with image-formation or the deflection of visual rays; they have the unique property of focusing light rays. In other words, burning mirrors have nothing to do with sight and everything to do with light. As we will see in the second part of this section, the earliest attempt to understand the focusing property of burning mirrors—in Euclid's *Catoptrics*—was as crude as it was misguided. Within roughly a century, however, when Diocles produced his exquisite study of spherical and parabolic burning mirrors, the analysis had achieved an extraordinary level of sophistication.

1. *The Rainbow*

Aristotle's study of the rainbow comes in the third book of the *Meteorology*, where he discusses certain atmospheric phenomena, such as lunar and solar halos, that are, by his account, due to reflection. However—and here we come to an interesting interpretive problem—it is not the reflection of light

but, rather, of visual flux (*opsis*) that creates such effects. What makes this point problematic, of course, is that the theory of sight Aristotle develops with such punctilio in *On the Soul* and *Sense and Sensibilia* (see Section I, pp. 25-27 above) is unequivocally intromissionist. He is in fact at pains there to refute the extramissionist alternative. Why this apparent contradiction? While there is no definitive answer to this question, it seems likely that Aristotle adopted the visual-flux model because it provided the standard framework for geometrical optics at the time (note, also, the use of a visual-flux model to explain shortsightedness in passage 1.7 above from the *Problems*, which, if not Aristotle's own work, is at least Aristotelian). Thus, Aristotle's use of the visual-ray model in the *Meteorology* probably reflects a methodological rather than a theoretical commitment. That said, let us turn to the following passage from *Meteorology*, III, 4, 373a35-373b25 (pp. 601-602):

[5.1] Sight is reflected from all smooth surfaces, such as are air and water among others. . . . But things are best reflected from water, and even in the process of formation it is a better mirror than air; for each of the particles, the union of which constitutes a raindrop, is necessarily a better mirror than mist. Now it is obvious and has already been stated [in III, 2, 372a29-372b6] that a mirror of this kind renders the colour of an object only, but not its shape. Hence it follows that when it is on the point of raining and the air in the clouds is in the process of forming into raindrops . . . if the sun is opposite, or any other object bright enough to make the cloud a mirror and cause the sight to be reflected to the object, then the reflection must render the colour of the object without its shape.

From this passage it is clear that Aristotle has in mind three basic prerequisites for the formation of a rainbow: a raincloud aborning through the coalescence of water-particles into droplets; a bright object (such as the sun or moon) directly opposite the cloud, and a viewer facing the cloud so that his sight can reflect to the bright object—in other words, so that the object is to his back. With these points in mind, Aristotle goes on in *Meteorology*, III, 4, 374b9-375a8 (p. 603) to explain the color-bands that can be distinguished within the rainbow.

[5.2] We must recognize . . . that white colour on a black surface . . . gives red; second, that sight when strained to a distance becomes weaker and less; third, that black is a sort the negation of sight: an object appears black because sight fails; so everything at a distance looks blacker, because sight does not reach it Clearly, then, when sight is reflected it is weakened and, as it makes dark look darker, so it makes white look less white, changing it and bringing it nearer to black. When the sight is relatively strong the change is to red; the next stage is green, and a further degree of weakness gives violet. . . . Hence also the rainbow appears with three colours [so that]

the outer band of the primary rainbow is red; for the largest band reflects most sight to the sun, and the outer band is largest. The middle band and the third go on the same principle. . . . The appearance of yellow [between the outer red and middle green bands] is due to contrast; for the red is whitened by its juxtaposition with green.

How, then, do we apply this analysis of color-perception to the model of reflection from individual droplets described in passage 5.1 above? For that we turn to *Meteorology*, III, 5, 375b16-376b21, where Aristotle describes a semicircle with its base GP on the horizon, its center K at the center of sight, and the sun rising at point G . Accordingly, as Aristotle describes it in *Meteorology*, III, 5, 375b30-376b15 (pp. 604-605):

[5.3] First, let the luminous body be rising at the point G [in figure 5.1], and let KM be reflected to G , and let the plane determined by the triangle GKM be produced. Then the section of the sphere will be a great circle. [Now] since both the points G and K and the line KG are given, the line MG will be given too; consequently, the ratio of the line MG to the line MK will be given too. . . . If, then, the semicircle. . . be revolved about the diameter GKP , the lines reflected from the points G, K at the point M will have the same ratio, and will make the angle KMG equal in every plane.²

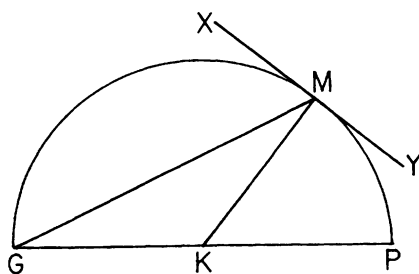


figure 5.1

Two things are worth noting about this explanation. First, the reflection that Aristotle describes is not bound by the law of equal angles, for ray KM , being a radius of the circle, strikes it orthogonally. Hence, the angle of incidence, KMY , is a right angle, which means that the angle of reflection GMX will be less than a right angle by the amount represented by angle KMG . Second, the position of point M on the semicircle, which defines the position of the rainbow, is unique, its location being a function of a particular ratio between GM and MK . The precise measure of that ratio is determined by finding two points, D and B , on diameter GP [in figure 5.2] such that $GD : DK = GB : BK = GM : MK$. Accordingly, M will lie at the intersection of semicircle GMP and the semicircle erected on segment DB of line GB (the so-called

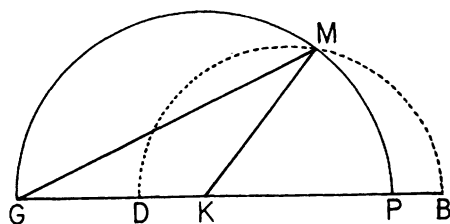


figure 5.2

Apollonian semicircle).³ The position of *M*, which defines the rainbow's arc, is therefore invariable with respect to *G* and *K*. So the curvature of the arc—which is a function of the size of the circle generated by rotating triangle *GKM* about axis *GKP*—is constant. The size of the arc can vary, however, depending on the position of *G* with respect to the horizon. Hence, as *G* rises above the horizon [in figure 5.3], diameter *GKP* of the semicircle tilts about *K* as a fulcrum, that tilt in turn pulling point *M* lower toward the horizon line until, at last, *G* rises high enough to make *M* disappear below the horizon-line.

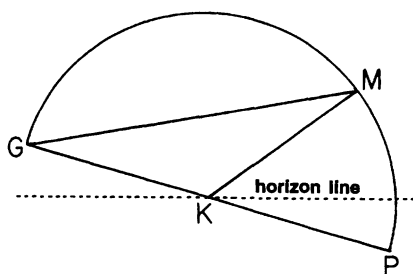


figure 5.3

The arc formed by the rotation of point *M* about axis *GKP* is a mere line, whereas the rainbow itself consists of several bands of color. The arc described by *M* therefore represents a mean about which those bands are grouped. In other words, the reflection involved in the rainbow's formation occurs within a range whose midpoint is defined by this arc. Now, as Aristotle claims in passage 5.2 above, the rainbow consists of three basic bands starting with red at the top and passing through green in the middle to violet at the bottom. If, then, the arc formed by *M* is taken to define the green band in the middle, the formation of the other two bands can be explained on the basis of the relative lengths of their generating ray-couples. For instance, the ray-couple forming the violet band below *M* is longer than ray-couple *KMG*, so the resulting visual impression along this ray-couple will be weaker.⁴ By the same token,

the ray-couple forming the red band above *M* is shorter than ray-couple *KMG*, so the resulting visual impression will be stronger. And, finally, as Aristotle asserts in passage 5.2 above, “the appearance of yellow [between the outer red and middle green bands] is due to contrast; for the red is whitened by its juxtaposition with green.”

Despite its meretricious exactitude and its superficial coherence, Aristotle’s explanation of the rainbow raises at least as many problems as it resolves. For one thing, the “law” of reflection underlying that explanation (i.e., the law defined by the specific ratio of *GM* to *MK*) flouts the law of equal angles. Why should reflection from small mirrors, as represented by individual raindrops, be any different from reflection from large mirrors? Moreover, that “law” is no law at all insofar as it is breached by the formation of the red and violet bands on either side of the arc defined by *M*. Another problem has to do with the formation of the secondary rainbow above the primary one. Two issues stand out in this case. First, there is a dark space between the upper band of the primary bow and the lower band of its secondary companion. According to Aristotle’s account of weakening by distance, however, that space should appear even lighter than the red band below it, since the distance comprehended by the ray-couple is even shorter. Second, the order of colors in the secondary bow is opposite that in the primary one—i.e., it runs from red at the bottom to violet at the top. Again, according to Aristotle’s account of weakening by distance, this should not happen, since the ray-couple yielding the impression of red, which is strongest, is longer than the ray-couple yielding the impression of green, which is weaker; and that ray-couple is in turn longer than the one yielding the impression of violet, which is weakest. Thus, the reversal of colors in the secondary bow is matched by a reversal of the weakening-principle described by Aristotle in passage 5.2 above.

2. *Burning Mirrors*

As we have already seen with Hero’s *Catoptrics*, and as we shall see in our selections below, there was a certain fascination during antiquity with mirrors and the effects that could be created with them. As far as burning mirrors go, we have the well-known testimony of Apuleius about Archimedes’ destruction of the Roman fleet at Syracuse using such mirrors. That testimony is quite late, however, and there is no contemporary mention of this astonishing event, an event that would surely have occupied the attention of the chroniclers of the day. Moreover, the technical difficulties of carrying out such a feat would have been insurmountable. As far as we know, the earliest surviving analysis of burning mirrors is to be found in the Euclidean *Catoptrics*, proposition 30 (pp. 340-42), which is reproduced below:

[5.4] [PROPOSITION 30] *Fire is ignited by [spherical] concave mirrors that are turned to face the sun.*

Let ABC [in figure 5.4] be a [spherical] concave mirror, EZ the sun, and T the mirror's center of curvature. From some point D let DT be drawn to centerpoint T and extended to B . Let ray DC also fall and be reflected to K . It will therefore reflect [to a point] above centerpoint T , for angle P formed at the [mirror's] surface is smaller than the remaining angle—i.e., BCD —formed at the [mirror's] surface. Let arc $AB = \text{arc } BC$, and from D let another ray, DA fall. It is therefore obvious that ray AD reflects to K , since arc $AB = \text{arc } BC$. In the same way we can demonstrate that all rays falling from D to the mirror and cutting off equal arcs intersect straight line BT at the same point [which lies] above T .

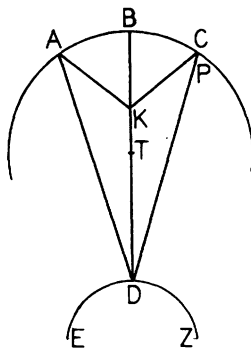


figure 5.4

Again, let ABC [in figure 5.5] be a [spherical] concave mirror and EZ the sun, and let ETB be drawn from some point E through center of curvature T . Let straight lines DTC and ZTA be drawn from other points [such as] D and Z . Accordingly, we have demonstrated earlier that rays emitted from E reflect back along themselves, because the angles formed at B are equal, insofar as they are formed by the diameter. Moreover, [the same holds for] rays emitted from Z , since the angles formed at A are equal, and [likewise for] rays emitted from D along DC , since the angles formed at C are equal. That all rays reflect back upon themselves is clear, for when they are extended through the centerpoint, they form semicircles, and the angles formed with semicircles are equal.

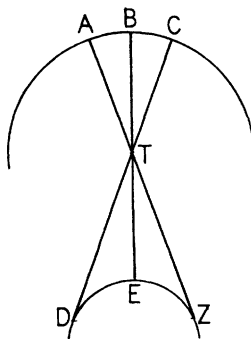


figure 5.5

Therefore, the reflections are made at equal angles, so [the rays] are reflected back upon themselves. Hence, all the rays emitted from every point will coincide with rays falling on the centerpoint. Heat is therefore concentrated at the center by these warming rays, so that tow will be ignited if it is placed at this point.

There is some dispute over whether this theorem is genuinely Euclidean or whether it represents a later, perhaps much later, and extremely maladroit interpolation. There is no gainsaying its superficiality. Moreover, it is incoherent, the first part establishing correctly that the center of curvature in a spherical mirror does not represent a focus for incident rays and the second

part contravening that very point. A far more sophisticated analysis of spherical burning mirrors is to be found in Diocles' *On Burning Mirrors*, which probably dates to the very late third century B.C. The following passage is from propositions 2 and 3, pp. 54-62, of that treatise:

[5.5] [PROPOSITION 2] Let AGM [in figure 5.6] be [a segment of] the circumference of a circle whose center is B , and let the circumference be met by some line AB . Let there be drawn from any two points on it (the circumference) two lines parallel to AB , namely LF and GS . Join BL and BG , and cut off arcs LM and GZ equal to arc LG . Then angle $BLM =$ angle BLG . And if we make angle $ZGX =$ angle LGS and angle $GLE =$ angle FLM , then angle $ELB =$ angle BLF , remainders (of angle $BLG -$ angle GLE and angle $BLM -$ angle FLM). But angle $FLB =$ angle LBE , [its] alternate, so angle $LBE =$

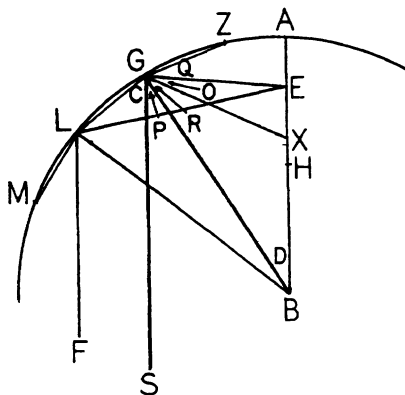


figure 5.6

angle ELB , so $BE = LE$. But $LE > EA$, because EA is the shortest of (all) the lines drawn from E to the circumference of the circle. So $BE > EA$, and also $BE > EG$, since $LE > EG$. So angle $O +$ angle $R >$ angle D ; and the angle alternate to angle D is angle P , so angle $O +$ angle $R >$ angle P . But the whole of angle $C +$ angle $P =$ angle $R +$ angle $O +$ angle Q . So angle $C >$ angle Q . But angle $Q +$ angle $O =$ angle C , so angle $R =$ angle P , remainders (of [angle $R +$ angle $O +$ angle Q] - [angle $O +$ angle Q] and [angle $C +$ angle P] - angle C). And angle $P =$ angle D , so angle $R =$ angle D , so $BX > XA$.

Then let AB be bisected at H . Hence it has been demonstrated that if we draw any number of parallel lines to AGM , the circumference of the circle, and they are reflected from circumference AGM so as to produce equal angles, then (the reflected lines) pass between points A and H , and no line among them passes between points B and H . The nearer one of the lines parallel to AB is to AB , the nearer to point H does its reflection (pass), and the farther

it falls on the area in which the burning takes place. Therefore whoever wants to make a burning-mirror from a section of a sphere need use only the surface produced by arc GA .

Diocles has thus demonstrated two important points about the use of spherical concave mirrors as burning mirrors, both points being somewhat unexpected. The first point is that, although there is no focal point for the reflected rays in the case of spherical concave mirrors, there is a fairly concentrated focal area for them. The second—and more surprising point—is that the burning capacity of such mirrors is not a function of the overall surface-area exposed to the sun: that is, bringing more of the sphere's surface into play does not increase the concentration of rays. The effective surface is therefore extremely limited.

It has long been supposed that, by the time Diocles wrote his treatise, it had been established that the proper shape for a true burning mirror is parabolic. Although there is no textual evidence to support the claim, it is generally assumed that Archimedes (d. 212 B.C.) was aware of that fact and, moreover, that his younger contemporary, Apollonius of Perga, was too. After all, Apollonius wrote the definitive treatise on conic sections. However, it is noteworthy that, although Apollonius established the focal properties of both the ellipse and the hyperbola, he had nothing to say on that score for the parabola. Whatever the case, the first known discussion of parabolic mirrors occurs in Diocles' *On Burning Mirrors*, of which proposition 1, pp. 44-48, is given in the following passage:

[5.6] [PROPOSITION 1] We believe that it is possible to make a burning instrument of glass such that it has a special property, namely that one can make lamps from it which produce fire in temples and at sacrifices and immolations, so that the fire is clearly seen to burn the sacrificial victims; this occurs, as we are informed, in certain remote cities, especially on the days of great celebrations: this causes the people of those cities to marvel.¹⁰ That is something which we too shall do.

Let there be a parabola KBM [in figure 5.8] with axis AZ , and let half the parameter of the squares on the ordinates be line BH .¹¹ Let BE on the axis be equal to BH , and let BE be bisected at point D . Let us draw a line tangent to the section at an arbitrary point, namely line QA , and

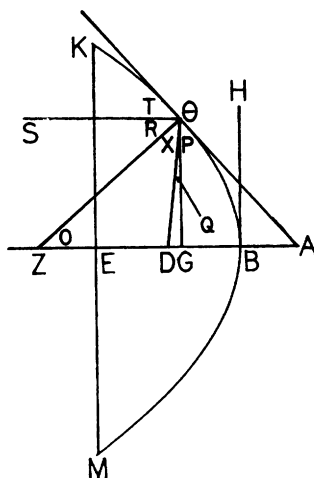


figure 5.8

draw line QG as ordinate to AZ . Then we know that $AB = BG$ ¹² and that the line drawn from Q perpendicular to QA meets AZ beyond E .¹³ So let us draw ZQ perpendicular to QA , and join QD . Then $GZ = BH$, and $BH = BE$, so $GZ = BE$. We subtract GE , common (to GZ and BE); then the remainder $GB = EZ$. But $GB = BA$, so $AB = EZ$. And $BD = DE$, because BE is bisected at D , so the sum $AD = DZ$. And because triangle AQZ is right-angled and its base AZ is bisected at D , $AD = DQ = DZ$.¹⁴ So angle $O = \text{angle } X$, and angle $A = \text{angle } P + \text{angle } Q$.¹⁵ So let a line parallel to AZ pass through Q , namely line QS . Then angle $O = \text{angle } R$, which is alternate to it, and angle $O = \text{angle } X$, so angle $X = \text{angle } R$ also. And angle $P + \text{angle } Q + \text{angle } X = \text{angle } R + \text{angle } T$, [both sums being equal respectively to] right angles, so angle $T = \text{angle } P + \text{angle } Q$, remainders. So when line SQ meets line AQ it is reflected to point D , forming equal angles $P + Q$ and T between itself and the tangent AQ . Hence, it has been shown that if one draws from any point on KBM a line tangent to the section and draws the line connecting the point of tangency with point D , e.g., line QD , and draws line SQ parallel to AZ , then in that case line SQ is reflected to point D , i.e., the line passing through point Q is reflected at equal angles from the tangent to the section. And all parallel lines from all points on KBM have the same property, so, since they make equal angles with the tangents, they go to point D .

Hence, if AZ is kept stationary, and KBM revolved (about it) until it returns to its original position, and a concave surface of brass is constructed on the surface described by KBM , and placed facing the sun, so that the sun's rays meet the concave surface, they will be reflected to point D , since they are parallel to each other. And the more the (reflecting) surface is increased, the greater will be the number of rays reflected to point D .

With this elegant demonstration, Diocles has established not only that a parabolic mirror focuses all incoming radiation to a true point but, equally important, that, unlike spherical mirrors, parabolic mirrors become increasingly effective as more of their surface-area is exposed to the light-source.

NOTES

¹In his commentary on Aristotle's *Meteorology*, Olympiodorus (late sixth century, A.D.) imputes to Ptolemy the opinion that the rainbow consists of seven colors. Unfortunately, we have no textual evidence whatever to support this claim.

²In other words, if we rotate the semicircle about axis GKP, point M will describe a semicircle, that semicircle representing the basic arc of the rainbow.

³For this explanation of how point M is located on the semicircle I depend heavily upon Carl B. Boyer, *The Rainbow: From Myth to Mathematics* (New York: Yoseloff, 1959). I am equally dependent upon him for the account of the formation of the rainbow's color-bands below.

⁴The closer to P the point of reflection approaches, the longer the resulting distance to G. Hence, while the length of the incident ray from K remains constant, the length of the reflected ray increases, so that the overall distance traveled between K and G increases.

⁵Since angle S = 30°, angle O = 30° by virtue of their equality. Angle S is also equal to the angle at G, because they are alternate angles between the two parallels TG and AB. Thus, the remaining angle E = 30°. Accordingly, angle O + angle E + angle TGB = a right angle, and thus so must angle R.

⁶That is, both are 60°.

⁷Since segment GN = 1/6 of the entire circumference of the circle, the chord of that segment, GN, is equal to the radius of the circle (by Euclid, *Elements* III, 15). Therefore, GK = 1/2 the circle's radius, and so does HA (by construction).

⁸Since BHZ is a 30°/60°/90° triangle, $BZ = 2ZH$. Thus $BZ^2 = 4HZ^2$. Thus, too, since $BZ^2 = BH^2 + HZ^2$, $BH^2 = BZ^2 - HZ^2 = 3HZ^2$. Thus, since $HZ^2 = 1/3BH^2$, $BZ^2 = BH^2 + 1/3BH^2 = 4/3BH^2$. Finally, since $BZ = BX$, then it follows that $BX^2 = 4/3BH^2$.

⁹Since $BX^2 = 4/3BH^2$, then $4BX^2 = 3BH^2$, so it follows that $2BX = 3BH$. Since $3 = 1.73+$, then half that quantity = .86+. Accordingly, $BX = 3/2 BH = .86BH$. Therefore, the difference between the two, HX, equals the remainder, i.e., .14. $.86/.14 = 6.14$. Thus, BH is a bit more than six times as large as BX.

¹⁰Note that the "practical" *raison d'être* for this effect—to arouse wonder and awe among religious celebrants—seems to be essentially the same as for the mirror-effect described by Hero of Alexandria in passage 3.35 above.

¹¹BH actually represents the distance from the vertex of the parabolic section to the vertex of its generating cone; since the parameter is twice that distance, then BH represents half the parameter.

¹²This equality stems from the fact that when the tangent from the ordinate cuts the parabola's axis, the section between that point of intersection and the parabola's vertex is equal to the section between the parabola's vertex and the point where the ordinate intersects the axis.

¹³It is a property of segment ZG formed by the intersections of the perpendicular from the tangent and the ordinate on the axis that it be equal to half the parameter, which is equal to BE (by construction). Thus, intersection-point Z must lie below point B , since intersection-point G lies below point B .

¹⁴Since triangle AQZ is right, then it can be inscribed into a circle with ZA as diameter and D as center. Hence, AD , DQ , and DZ , being radii, are all equal.

¹⁵This follows from the fact that triangles DQZ and AQD are isoceles.

CONCLUSION

The immediate course of mathematical optics after Ptolemy is difficult to trace with any accuracy because of a lack of textual evidence. All indications are, however, that the science of optics suffered considerable decline during late antiquity. There is fairly clear evidence, for instance, that, while Euclid's optical work was still in effective circulation (as witness Theon of Alexandria's fourth-century commentary on Euclid's *Optics*), Ptolemy's optical masterpiece seems to have sunk into near-immediate oblivion. True, an absence of texts is not proof of an absence of scholarly interest or activity; but, given the scant textual evidence available—including Anthemius of Tralles' rather maladroit sixth-century study of the focusing-property of parabolic mirrors—we must conclude that the period between roughly 200 and 850 A.D. was one of stagnation, if not outright regression.

The situation changed dramatically with the Arabs. No doubt because of the translation-program sponsored by al-Ma'mūn (fl. c. 820), various Greek optical texts came under serious scrutiny during the first half of the ninth century. Out of that scrutiny came one of the earliest optical treatises to be composed in Arabic, the *De aspectibus* of Ya'qūb al-Kindī (d. c. 870). While deeply rooted in the Euclidean optical tradition, this treatise nonetheless shows signs of genuine creativity, particularly in the analysis of radiation from point-sources. The model that al-Kindī constructs in proposition 14 is as follows: let circular segment ABD with centerpoint C [in figure 6.1] represent the surface of the cornea, and let points A, B, and D represent point-sources of visual radiation. Each point will therefore create a hemisphere of radiation, A's being represented by arc EF, B's by arc GH, and D's by arc KL. All of these hemispherical segments overlap toward point X, where radial line BC intersects arc EL of overall radiation. Thus, the more points between A and D that are taken into account, the more concentrated the area of overlap toward point X. "Therefore," al-Kindī concludes, "it has been demonstrated that the center [of the

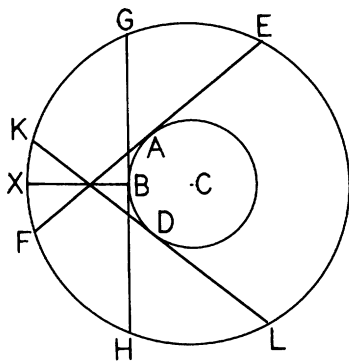


figure 6.1

hemisphere of radiation at point X] is intensely illuminated.. And what lies closer to that point is more intensely illuminated than what lies farther from it. For more light falls upon that point—i.e., it is illuminated by more point-sources of light.”

Within a century of al-Kindi's death (c. 870), mathematical optics had reached such a high analytic level that the mathematician Ibn Sahl (fl. c. 980) was able not only to derive the sine-law of refraction but to use it in demonstrating the focal properties of hyperbolic lenses. In this he anticipated Kepler (d. 1630) and Descartes (d. 1650) by some six-and-a-half centuries. But far more historically significant than Ibn Sahl was his younger contemporary, Ibn al-Haytham (Latinized as “Alhacen”), author of a monumental optical synthesis entitled *Kitāb al-Manāẓir* (“Book of Optics”). Closely modeled after Ptolemy's *Optics*, this work brought together several analytic strands, including Galenic ocular anatomy and physiology, Euclidean-Ptolemaic ray-analysis, experimentalism, and Aristotelian physics. The result is a true integration that represents more than the sum total of its separate parts.

In a real sense, then, Alhacen's *Kitāb al-Manāẓir* marks the culmination of the Greek—particularly the Ptolemaic—optical tradition. Yet in an equally real sense it marks the beginning of a new and fundamentally antithetical optical tradition. For, whereas Alhacen's predecessors seem to have been unanimous (or at least virtually so) in supporting the Euclidean-Ptolemaic model of visual radiation, Alhacen was adamant in rejecting it. Instead, he proposed an intromissionist alternative based on the radiation of light and color from external objects into the eye. Unlike Euclid and Ptolemy, moreover, Alhacen made light a direct agent in vision. No longer a mere prerequisite for visibility, light was now visible in its own right. As such, it was functionally independent of the visual process and, therefore, a subject for physical analysis in its own right.

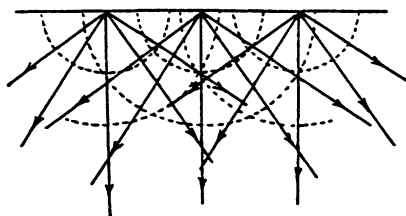


figure 6.2

By rendering light per se visible, Alhacen pointed the way toward the development of modern physical optics, yet, despite such intimations of modernity, Alhacen's work was still firmly entrenched in the classical tradition of ray-analysis with its emphasis on sight rather than light. A look at his basic

theory of visual perception will make this point clear. According to Alhacen, every point on a luminous or illuminated object radiates its form in all possible directions, thus creating a hemisphere of radiation (note the similarity to al-Kindi's model of punctiform radiation). Each such hemisphere can be resolved into innumerable individual rays [as represented in figure 6.2]. Accordingly, every point on the luminous or illuminated object is radially connected to every point facing it on the surface of the eye [see figure 6.3].

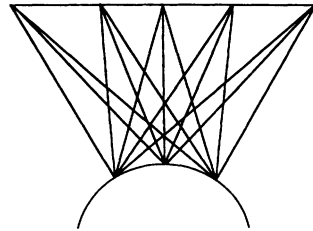


figure 6.3

As a receptor of such point-forms, the eye has a peculiar structure adapted its visual function. The ocular globe itself [as represented in figure 6.4] is contained within an outer sheath that consists of the sclera and its corneal extension at the front. Its centerpoint is at D. Inside this globe is a smaller globe called the uvea, whose centerpoint C lies in front of D. The opening at the front of the uveal sphere forms the pupil. Next in line comes the crystalline lens. Its anterior surface is a segment of a sphere with its center of curvature at D, and its posterior surface is a segment of a more gently curved sphere. Finally, at the rear of the uveal sphere is an opening for the hollow optic nerve that originates at the front of the brain. This nerve conveys visual spirit from the

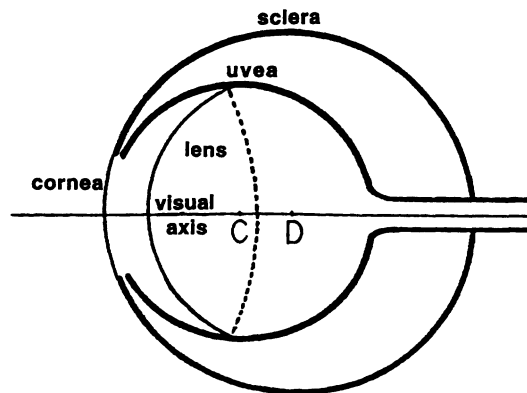


figure 6.4

brain all the way to the anterior surface of the crystalline lens, which is thereby rendered visually sensitive. All of the spaces within the eye are filled with optic humors of different refractivity. The region between the cornea and the anterior surface of the crystalline lens is replete with albugineous humor, which is more refractive than air. The crystalline lens, in turn, is filled with glacial

humor, which is more refractive than the albugineous humor in front of it. And the rest of the uveal sphere behind the lens is filled with vitreous humor, which is even more refractive than the glacial humor in front of it.

So constituted, the eye--and, more specifically, the crystalline lens--receives the point-forms of light and illuminated color that are radiated from all the point-sources on facing visible objects. The ensuing impression should therefore be as chaotic as possible. We should not, in short, see defined images. That we do, Alhacen concludes, is because of the peculiar visual sensitivity of the crystalline lens that lets it accept, or "feel," only those forms that strike it orthogonally. The rest it ignores because of their relative weakness, so what is effectively impressed on the lens is a point-by-point representation of the original visible surface. The lens's refractive disposition also comes into

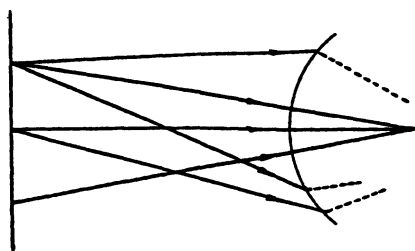


figure 6.5

play by allowing those rays that strike it orthogonally to pass through unrefracted while deflecting the rest out of consideration [see figure 6.5]. What results is a cone of effective radiation whose base is formed by the visible surface and whose vertex lies at D, the center of the ocular globe. Before actually

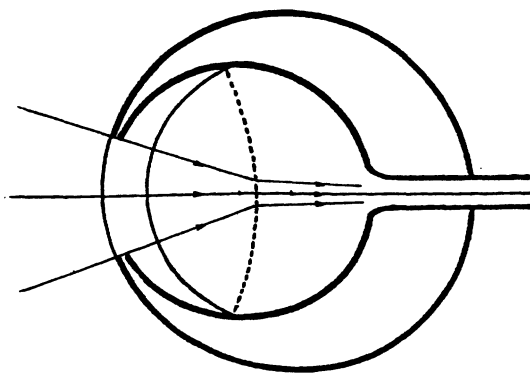


figure 6.6

intersecting at D, however, the rays-cum-image are refracted at the posterior surface of the lens so as to be funneled in proper, upright order along the visual axis into the hollow optic nerve whence it eventually passes into the brain for perceptual judgment [see figure 6.6].

From this account it should be clear that Ptolemy's cone of visual radiation and Ibn al-Haytham's cone of visible radiation are perfectly congruent; that is, every Euclidean-Ptolemaic visual ray extending outward from the center of sight to the visible object has an Alhacenian counterpart extending inward from the visible object to the center of sight. As far as ray-analysis is concerned, then, the two accounts are mathematically equivalent because they share the same analytic perspective—from the center of sight outward. This is not, however, to deny that Alhacen made significant advances over Ptolemy. At the level of anatomy and physiology, for instance, Alhacen's account is far more comprehensive and profound than Ptolemy's, and the same holds at the level of perceptual psychology. At the level of mathematics, moreover, Alhacen's analysis of specific optical problems, particularly in reflection and refraction, represents a quantum leap in sophistication over Ptolemy's.

Small wonder, then, that Alhacen's optical synthesis gained canonical status within Scholastic circles so soon after being translated into Latin (c. 1200). Indeed, the logical force of his account was such that, within only a few decades of its first appearance in Latin form, it had spawned a handful of crucial derivative works, including Roger Bacon's *De multiplicatione specierum* (c. 1260) and *Perspectiva* (c. 1265), Witelo's *Perspectiva* (c. 1275), and John Pecham's *Perspectiva communis* (c. 1280). Along with the Latin version of Alhacen's own treatise, under the title *De aspectibus*, these works formed the core of the so-called Perspectivist tradition, which flourished until the early seventeenth century. Within this tradition the science of optics achieved a remarkably high technical level during the Middle Ages and Renaissance.

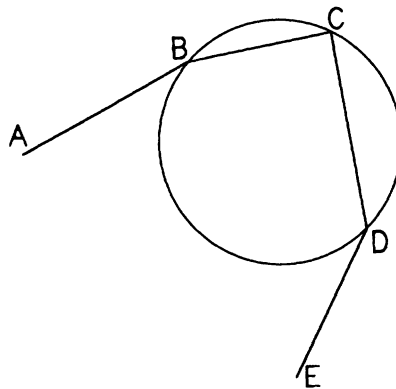


figure 6.7

A salient example is Theodoric of Freiburg's early-fourteenth-century resolution of the rainbow-problem according to a combination of refractions and reflections within each raindrop. Let sphere **BCD** [in figure 6.7] represent a raindrop, and let **AB** represent a solar ray. Let **AB** strike the raindrop at such an angle that it passes through rather than reflecting from interface **BD**. Accordingly, when **AB** enters the optically denser raindrop, it will be refracted toward the normal along path **BC**. If, in turn, the angle of incidence at point **C** is adequate to cause reflection, the ray will be diverted to point **D**, where it will re-enter the air to be refracted away from the normal along **DE**. On reaching the eye, **DE** will therefore yield a specific color-impression that depends ultimately upon the original angle of incidence at the raindrop's surface. As we now know, in fact, the color we see along **DE** is due to the prismatic effect created by the raindrop, and that effect is, as Theodoric surmised, a function of the original angle of incidence at the raindrop's surface. Theodoric was therefore essentially correct in his analysis of the rainbow, at least at the descriptive level. And he was not alone. Working independently at almost precisely the same time as Theodoric, an Arab disciple of Alhacen, Kamāl al-Dīn, arrived at virtually the same explanation of the rainbow.

Despite the weight of authority that it accrued between roughly 1260 and 1600, the Perspectivist account of vision did not so overwhelm its Euclidean-Ptolemaic forebear as to render it moot. On the contrary, several Scholastic thinkers, Roger Bacon foremost among them, viewed the two accounts as complementary rather than contradictory, and some later thinkers, such as the sixteenth-century optical writer, Francesco Maurolyco, felt free to shift between intromissionist and extramissionist frameworks when it suited their analytic purposes (cf. the discussion of Aristotle on p.151 above). Indeed, the relatively rich Latin textual traditions of Euclid's *Optics* (34 manuscript copies) and *Catoptrics* (52 manuscript copies), as well as of Ptolemy's *Optics* (13 manuscript copies), bespeak a continuing and lively interest in those works right up to the end of the sixteenth century. And why not? Since both the Perspectivist and Euclidean-Ptolemaic accounts share the same analytic framework—in terms of a cone of radiation with its vertex at the center of sight—then, as far as ray-analysis is concerned, Perspectivist optics can be viewed as a mere elaboration on, not a refutation of, its Euclidean-Ptolemaic counterpart.

Thus linked through analytic focus, both the Ptolemaic and Perspectivist accounts shared the same fate in the early seventeenth century when Kepler subjected the latter to a devastating critique in his *Ad Vitellionem paralipomena* ("Emendations to Witelo") of 1604. The focus of his critique was the cone of radiation. Why, Kepler mused, should the rays that form the visual image be limited to those striking the lens orthogonally? Why, in other words, should the crystalline lens be exempt from the laws that govern all other transparent

bodies of equivalent kind—for example, glass flasks filled with water? Kepler's answer was that it should not. On that basis he transformed the lens from a sensitive form-selector to a mere focusing-device. Accordingly, as figure 6.8

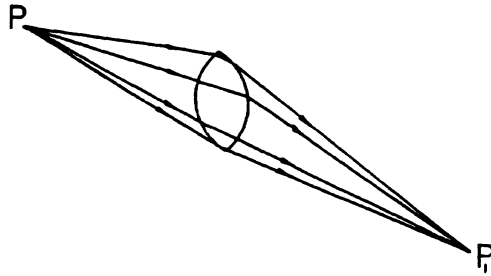


figure 6.8

illustrates, the lens is understood to accept *all* of the rays reaching it from any given point-source P , bringing them to focus in a point-image P_1 whose location is determined by the lens's curvature and refractivity. The more point-sources that are taken into account, the more point images are projected, until, finally, an inverted point-by-point representation of the original object is

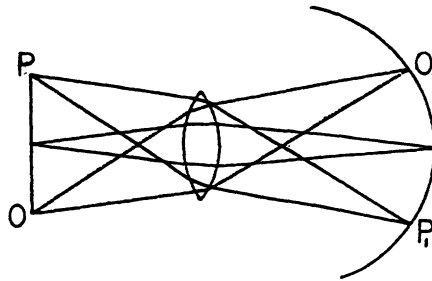


figure 6.9

formed behind the lens [as in figure 6.9]. By Kepler's reasoning, then, the eye is nothing more than a camera within which the images of external objects are cast by the lens upon the retinal screen at the back.

By "objectifying" the account of sight in this way, Kepler did away in one fell swoop with the central analytic conventions of both the Euclidean-Ptolemaic and the Perspectivist accounts of vision. First and foremost, the eye or crystalline lens was no longer viewed as a sensitive agent in the visual process; it was reduced to absolute passivity. Second, the cone of radiation was now functionally irrelevant because all of the rays reaching the eye from a given object, not just the orthogonals, go into the formation of visual images. This meant, finally, that the center of sight no longer served any analytic purpose,

because it no longer represented the intersection-point—either actually or potentially—for the visible radiation transmitted by the lens. Thus bereft of its traditional analytic viewpoint, optics was at last ready to make the transition from science of sight to science of light, a transition that was completed by the end of the seventeenth century.

GENERAL INDEX

- Academy 18
aisthesis koine 39
aktis 76n
albugineous humor 163, 164
Alcmaeon 27
Alexandria 16-18, 38
Alhacen 162-166
Almagest 11, 17, 140
anomaly 72, 101
Anthemius of Tralles 161
Apex 37, 72, 77n
Apollonian semicircle 152
Apollonius of Perga 157
Apuleius 153
aqueous humor 36
Archimedes 153, 157
Aristotle 11, 12, 14, 15, 17-19, 21n, 24-28, 30-34, 38-40, 48n, 147n, 149-153, 159n, 162, 166
astrolabe 145n
atoms 18
atomists 24, 26
axis: *See* common axis, proper axis, visual axis.
- Bacon, Roger 165, 166
Bevilacqua, Fabio 8
brain 35-37, 46, 48n, 163, 165
Boyer, Carl B. 159n
Burton, H. E. 16, 21n
- cause: *See* formal cause, material cause.
choroid 36
Circle of Willis 48n
Chrysippus 18
Cohen, M. R. 8 16, 17, 21n
color 12, 14, 18, 19, 25-27, 29-34, 36-45, 48n, 62, 100, 150-153, 159n, 162, 164, 166
comets 149
common axis 72-74
common sensibles 39, 40, 46, 51, 62
common sensibility 39, 40
cone: *See* visual cone.
cornea 25, 34, 36, 37, 48n, 76n, 86, 87, 161, 163
crystalline humor 36, 37. *See also* glacial humor.

crystalline lens 36, 37, 163-167

Democritus 13-15, 25, 26, 34

density: *See* optical density.

Descartes, René 162

Diocles 12, 18, 149, 155, 157, 158

Diogenes Laertius 13-15

Diogenes of Apollonia 35

dioptr 82, 123n

diplopia 69, 71, 72

directional privilege of rays 55, 56

Drabkin, I. E. 8, 16, 17, 21n

eidola 24, 25

Empedocles 27, 28, 34

emphasis 25, 34, 35, 87

Epicurus 18, 24, 25

equal-angles law: *See* law of equal angles in reflection.

ether 134, 135, 137, 146n, 147n

Euclid 8, 9, 11, 13-17, 21n, 24, 28, 30, 48n, 51, 52, 55-58, 62, 67, 69, 74, 76n, 80, 83-87, 94, 96, 97, 101, 105-108, 110, 111, 119, 123n, 124n, 125n, 126n, 127, 129, 149, 153, 154, 159, 161, 162, 165-167.

Eugene of Sicily 17

first differences 145n

formal cause 26

Galen 12, 14, 18, 29, 30, 34-38, 51, 74, 76n, 162

ghost image 99, 100

glacial humor 163, 164. *See also* crystalline humor.

Governing Faculty 37, 63, 64, 76n

Ibn al-Haytham 162. *See also* Alhacen

hegemonikon 37. *See also* ruling principle

Heiberg, J. L. 16, 17, 21n

Hero of Alexandria 11, 13, 15-17, 28, 54, 55, 79-81, 85, 96, 120-122, 123n, 127, 153, 159n

humor: *See* optic humor.

Ianniello, Maria Grazia 9

illusionism in painting 46

illusion: *See* oculogyral illusion, visual illusion

image: *See* ghost image, real image, virtual image

incidental sensibles 39, 40, 46

iris 36

Kamāl al-Dīn 166

Kepler, Johan 76n, 162, 166, 167

al-Kindī 161-163

Knorr, Wilbur 21n

law of equal angles in reflection 40, 55, 79-83, 94, 151, 153

Lejeune, Albert 9, 17, 21n

light 8, 11, 12, 14, 25-34, 41, 42, 48n, 53, 54, 76n, 149, 158, 162, 164, 168

Lindberg, David C.. 46

Lucretius 12, 18, 24, 25, 49n

Lyceum 17, 19

al-Ma'mūn 161

material cause 26, 35

Maurolyco, Francesco 166

Moerbeke, William of 16

Moon Illusion 140

myopia 28, 48n. *See also* shortsightedness.

oculogyral illusion 43, 46

Olympiodorus 159n

opsis 48n, 76n, 150

optic chiasma 36, 37, 77n,

optic humor: *See* albugineous humor, aqueous humor, crystalline humor,
glacial humor, vitreous humor.

optic nerve 29, 35-37, 163, 165

optical density 127, 133-135, 140, 144n

passion: *See* visual passion.

Pecham, John 165

Plato 11, 14, 17, 18, 24, 28-32

pneuma 29, 30, 35-38. *See also* psychic pneuma, vital pneuma.

principium 37, 77n

principle of reciprocity in refraction 144n

proper axis 73, 74, 77n. *See also* visual axis

proper sensibles 31, 38, 41. *See also* special sensibles.

psychic pneuma 35-37

pupil 25, 28, 35-37, 48n, 51, 52, 74, 76n, 95, 163. *See also* cornea.

Pythagoreans 18, 27

real image 79, 98, 100, 110, 111, 113, 115, 117, 125n

rete mirabile 48n. *See also* retiform plexus.

retina 36, 167. *See also* retiform tunic.

retiform plexus 36, 48n

retiform tunic 36

ruling principle 37, 76n. *See also* hegemonikon.

Ibn Sahl 162

Schmidt, Wilhelm 17, 21n

sclera 36, 163

seal-and-wax analogy 25-27

second differences 145n

- secondary rainbow 153
 sensibles: *See* common sensibles, incidental sensibles, proper sensibles, special sensibles.
 shortsightedness 28, 150. *See also* myopia.
 sine-law of refraction 7, 8, 145n, 162
 special sensibles 39. *See also* proper sensibles
 Stoics 14, 18, 24, 29-32, 37
 surface of reflection 95
- Takahashi, Ken'ichi 21n
 Theodoric of Freiburg 166
 Theon of Alexandria 16, 17, 161
 Theophrastus 12, 14, 19, 25, 28, 34, 35
 Thrasyllus 13
 tunics of the eye: *See* choroid, retiform tunic, retina, sclera, uvea.
- uvea 163, 164
- ventricles of the brain 35, 36
 Ver Eecke, Paul 16, 17, 21n
 virtual image 79, 97, 98, 110, 111, 113, 115, 117, 125n
virtus regitiva 37. *See also* Governing Faculty
 visual angle 56-58, 61, 67, 76n, 108, 114, 141, 142
 visual axis 53, 56, 66, 69-73, 76n, 77n, 86, 165. *See also* common axis, proper axis.
 visual cone 51-56, 63-65, 69-74, 76n, 86, 87, 95
 visual illusion 40, 43-46, 66. *See also* oculogyral illusion
 visual passion 32
 visual spirit 43, 163
visus 48n
 vital pneuma 36
 vitreous humor 36, 164
- Witelo 165, 166